



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

SIMULACIÓN NUMÉRICA DE PROPAGACIÓN LINEAL
Y NO LINEAL DE ONDAS ACÚSTICAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

P R E S E N T A :

ROBERTO VELASCO SEGURA

DIRECTOR DE TESIS:

DR. PABLO LUIS RENDÓN GARRIDO

2007



FACULTAD DE CIENCIAS
UNAM



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Simulación numérica de propagación lineal y no lineal de ondas acústicas

Roberto Velasco Segura

Director de tesis: Dr. Pablo Luis Rendón Garrido

*A mi madre,
a mi hermana.*

Agradecimientos.

Con el perdón de mis muy atentos, agudos y bien intencionados correctores me voy a permitir enredarme un tantito. Ocasión que puede usted aprovechar para darse una idea del terrible destino que hubiera tenido este texto de no haber pasado por las manos de ellos, lo cual yo aprecio y agradezco profundamente.

Todos ustedes que ahora lo leen creo que lo saben y lo sabían, ha sido un largo camino y nos han sucedido muchas cosas, juntos y separados. Para mi también este camino ha estado lleno de sorpresas de todo tipo, un montón de cosas que se han mezclado, donde he recibido el apoyo en todas las formas que podía imaginarlo y a decir verdad también de personas que no imaginaba. Por estas razones es que no puedo evitar en esta ocasión sentir agradecimiento hacia muchas personas que a decir verdad han tenido poco que ver en la manufactura de este texto, pero han tenido papeles fundamentales en todas esas experiencias que de alguna manera culminan ahora (más espero que no terminen). Por eso me da gusto agradecer a palabra escrita ahora a todos ustedes, y sin duda, muy en contra de mi voluntad he tenido la torpeza de omitir alguno.

En primer lugar a mi padre, mi hermano, mi madre y mi hermana, a todos ustedes gracias porque me han aguantado comprensivamente, me han apoyado incluso cuando no estamos de acuerdo, me han educado, me han detenido cuando era necesario, me han querido todos estos años como yo a ustedes.

A mi escuela, la gran Universidad Nacional Autónoma de México, y desde luego en particular a la Facultad de Ciencias, y más en particular a sus maestros que han compartido desinteresadamente sus instrucciones pero también sus ideas, a veces sus enfados, sus desconciertos y sus puntos de vista.

A mi grupo de sinodales, los Doctores Federico Sabina Ciscar, Catalina Stern Forgach, Felipe Orduña Bustamante y Xavier Hernández Döring, que paciente y amablemente han discutido y señalado en mi trabajo una buena cantidad de aspectos a corregir (que espero haber atendido suficiente). Desde luego de modo muy especial a Dr. Pablo Rendón Garrido por la infinita paciencia que has tenido conmigo, por todo el tiempo que has dedicado a este proyecto, por todo el apoyo que he recibido de tu parte, por tu amistad, gracias. De modo muy especial también a la Dra. Martha Rosete Aguilar, por tu amistad y por el profundo apoyo que me has brindado en muchas ocasiones.

A una lagartija: 'por estar a mi lado y todo lo que vivimos.

A todos lo humanoides por escucharme, por compartir su crecimiento conmigo, por escuchar las incoherencias que se me caen a plena madrugada (o a media tarde). También por su paciencia. A Humano por el espacio, y por la oportunidad de buscar el imposible equilibrio entre el ser y no estar.

A la familia Centeno por ayudarme con mis crisis nerviosas, por su amistad, gracias.

A la toda la familia Basauri, a la Sra. Susana y el Sr. Manolo, por el cariño que compartimos y todo el apoyo que me han dado, a Manuel porque aunque no te entiendo no puedo dejar de creer en ti, en que eres y seguirás siendo grande para nosotros. A Adriana por levantar mi cabeza y cultivar amistad sincera.

Al grupo II de México, no solamente por lo bien que la pasamos muchas veces, sino también por haber incidido en mi formación de modo constante desde que yo tenía escasos 11 años.

A Toñito, con todo cariño, que has sido padrino, amigo, compañero, sastre, doble seis y mucho más. Estoy seguro de que no entiendo la mitad de lo que dices, pero no dejo de intentarlo.

A Eduardo González, porque más de una vez has estado ahí para levantarme, porque has compartido la conmigo

la música, los buenos y los malos consejos.

Al Sr. Jorge, porque lo admiro y por tolerarme todas las veces que dice cosas que no entiendo, o no las dice, por abrirme la puerta de su casa y su confianza, gracias.

A la Sra. Patricia, Eduardo y el Sr. Eduardo, por brindarme su confianza y su cariño.

A mi maestro Emanuel Montes por ayudarme descubrir lo que hay escondido en la expresión de un rastro. A Daría y Pepe por dejarme entrar su casa y tocar ese maravilloso mundo del corazón expuesto.

A mi maestra Brigitte Loire, porque tu presencia es un ejemplo, aunque no me gustara hacer manchones, y aunque no hubiera aprendido nada, me siento agradecido de haber estado contigo.

A la familia Uribe, a Geronimo, Adriana, Paco, Abigail, Gabriel, porque a cada uno los admiro de verdad y los estimo como amigos que somos.

A Venus, La Negrita, La Trapita, Blues, Tequila.

A la Dona, gracias, no olvido.

A mis todos mis tíos, David, Hector, Roberto, Victoria, Araceli, Salvador, Araceli, Maricruz, Jaime, Esmeralda y mis primos Hector, Jaime, David, Salvador, Ayari, Araceli, Sofía, Michel, Vanya, Carla, Angélica, Tania, Diané, Zoé, Mauricio, Darinka, Sofía, Ximena porque somos familia, para lo que venga.

A Tábata por rayar mis cuadernos, a Judith, a Tania y a toda la familia por el apoyo y la amistad que nos han dado desde siempre.

A María Luisa, porque, sirva la ocasión para decirlo, has abierto muchas veces puertas en mis ojos, en mis oídos. Eres como un colibrí que un día se aparece y me cambia los zapatos por patines.

A Sophie, aunque nos duela leerlo.

Más, sin en cambio, también a Alvaro y Karlita, pues...

A la familia Pruneda y en especial a Laura por haberme escuchado, por habernos escuchado.

A todo el equipo de Caiinsa, y en particular a Gerardo y a Guillermo, por ofrecerme su conocimiento, que no se si lo parezca pero ha sido muy importante en este trabajo, gracias por su amistad y su paciencia.

A mis compañeros y amigos de la escuela (bueno aquí estaba una larga larga lista de nombres pero mejor digamos a toda la banda de ciencias) por los buenos ratos que pasamos en clases y fuera de ellas.

A Cuauhtémoc y su alma de tahúr, a Rafael, Susana, Ernesto, Iván, David, Miguel, Daniela, Ricardo, Mauricio, Karina, Hector, Eduardo, la Sra. Leticia, Miguel y Lazlo (que hizo todo el trabajo rudo), por su amistad y sacar a flote todo lo que ha llegado.

Al proyecto PAPIT IN116205, cuyo responsable es el Dr. Pablo Luis Rendón Garrido. Por el apoyo económico.

Índice

1. Introducción.	7
2. Ondas.	9
2.1. Derivación de la ecuación de onda lineal.	10
2.2. Caso unidimensional: ecuación de advección y solución de D'Alembert.	14
2.2.1. Otra manera de expresar la ecuación de onda: ecuación de advección.	15
2.2.2. Condiciones iniciales.	16
2.2.3. Curvas características.	18
2.3. Principio de superposición.	20
2.4. Dispersión y difusión.	21
3. Ondas de choque.	23
3.1. Ecuación de Burgers no viscosa.	23
3.2. Condiciones de Rankine-Hugoniot.	26
3.3. Problema de Riemann.	28
4. Conceptos básicos de los esquemas numéricos de diferencias finitas.	32
4.1. Aproximaciones a las derivadas.	33
4.2. Discretización.	34
4.3. Esbozo de esquemas numéricos de diferencias finitas.	36
4.4. Esquemas explícitos e implícitos.	37
4.5. Número CFL.	38
4.6. Notación polinomial.	39
4.7. Notación matricial.	40
4.8. Monotonicidad.	41
4.9. Expresiones derivadas de las ecuaciones diferenciales.	41
4.10. Error de truncamiento local, consistencia y orden de un esquema numérico.	42
4.11. Convergencia y estabilidad.	44
4.12. Condiciones iniciales.	47
4.12.1. Condiciones iniciales periódicas.	47
4.12.2. Condiciones iniciales de soporte compacto.	48
4.12.3. Equivalencia entre ambos tipos de condiciones iniciales.	50
4.13. Causalidad numérica.	51
4.14. Dispersión numérica y difusión numérica.	54
5. Esquemas numéricos relativos a la ecuación de advección.	55
5.1. Esquema contracorriente.	55
5.2. Esquema de Lax-Friedrichs.	59
5.3. Esquema de Lax-Wendroff.	62
5.4. Esquema de Wendroff.	66
5.5. Implementación.	70
5.5.1. Parámetros de control, fijos y de evaluación.	70
5.5.2. Implementación de los esquemas explícitos: contracorriente, Lax-Friedrichs y Lax-Wendroff.	73
5.5.3. Implementación del esquema implícito: Wendroff.	74
5.6. Resultados.	76
5.6.1. Gráficas de propagación.	79
5.6.2. Amplitud relativa con condiciones iniciales periódicas.	83
5.6.3. Desplazamiento con condiciones iniciales periódicas.	84
5.6.4. Norma con condiciones iniciales periódicas.	86

5.6.5.	Condiciones iniciales de soporte compacto.	88
5.6.6.	Resumen de observaciones.	90
6.	Ecuaciones modificadas para esquemas numéricos de la ecuación de advección.	92
6.1.	Solución de la ecuación de advección-difusión	93
6.2.	Solución de la ecuación de advección con dispersión.	94
6.3.	Expresiones derivadas de las ecuaciones diferenciales.	96
6.4.	Valor de α para el esquema contracorriente	96
6.5.	Valor de α para el esquema de Lax-Friedrichs.	97
6.6.	Valor de μ para el esquema Lax-Wendroff.	98
6.7.	Valor de μ para el esquema de Wendroff.	99
6.8.	Implementación y comparación con resultados numéricos.	101
6.9.	Otras ecuaciones modificadas.	107
7.	Esquemas numéricos relativos a la ecuación de Burgers no viscosa.	108
7.1.	Esquema primero.	108
7.2.	Esquema segundo.	110
7.3.	Esquema tercero.	111
7.4.	Implementación	113
7.4.1.	Esquemas explícitos: primero y tercero.	114
7.4.2.	Esquema implícito: segundo.	114
7.5.	Resultados	115
8.	Conclusiones y propuestas para trabajos posteriores	121
A.	Material digital interactivo.	125

1. Introducción.

Uno de los retos que frecuentemente aparece en diversos problemas de física consiste en encontrar soluciones exactas a diferentes tipos de ecuaciones diferenciales, tanto ordinarias como parciales. Aunque este campo ha sido ampliamente estudiado y son conocidas un buen número de herramientas para encontrar dichas soluciones, algunos casos, que frecuentemente son del tipo no lineal, sobrepasan los alcances de dichas herramientas y para ellos se conocen solamente algunas soluciones bajo condiciones especiales o se desconocen por completo las soluciones. Los métodos numéricos constituyen una alternativa en estos casos. A grandes rasgos éstos son modelos algebraicos, relacionados de alguna manera con la ecuación diferencial que modelan, y que ofrecen una solución *numérica* que aproxima la solución exacta de la ecuación diferencial original. Dado el carácter algebraico particular de los métodos numéricos siempre es posible encontrar su solución y es esta la ventaja que ofrecen, particularmente en los casos mencionados donde la solución exacta de la ecuación diferencial es desconocida.

Si bien los métodos numéricos son una herramienta aplicable para diferentes fines, también constituyen un campo de estudio en sí mismos. Esto se debe en primer lugar a que para su aplicación es necesario conocer la calidad de la aproximación de la solución numérica con respecto a la solución exacta de la ecuación diferencial, y saber si existen parámetros de los cuales dependa dicha calidad. En segundo lugar se debe a que existe una gran variedad de métodos numéricos cuya calidad de aproximación es diferente. Para el estudio de las propiedades y limitaciones de los diferentes métodos numéricos una estrategia común es utilizarlos con ecuaciones diferenciales cuyas soluciones exactas son conocidas, de manera que se puedan comparar los resultados numéricos con las soluciones exactas y así hacer una valoración de la calidad de los resultados numéricos; es de esta manera como se ha procedido en el presente texto.

En particular dentro del área de acústica aparecen múltiples ecuaciones diferenciales cuya solución es relevante para diversas aplicaciones. Con el fin de focalizar el estudio en el presente texto se ha escogido el fenómeno de propagación de perturbaciones en medios homogéneos. A dichos casos corresponden ecuaciones diferenciales parciales del tipo hiperbólico, de las cuales se han escogido para su estudio dos casos sencillos: la ecuación de advección y la ecuación de Burgers no viscosa. La primera de estas ecuaciones es una versión monodireccional de la bien conocida ecuación de onda lineal y la segunda tal vez constituye el caso más sencillo de propagación no lineal. Dentro del presente trabajo se implementan y estudian siete esquemas numéricos del tipo de *diferencias finitas* correspondientes a estas dos ecuaciones, cuatro esquemas para la ecuación de advección y tres para la ecuación de Burgers no viscosa. Dichos esquemas han sido escogidos buscando que sean representativos de diferentes tipos de esquemas, los cuales serán caracterizados más adelante.

A partir de los años cincuentas ha habido importantes contribuciones teóricas con respecto a los métodos numéricos, que han permitido extender ampliamente el espectro de aplicaciones y también su campo de estudio. El desarrollo de esta disciplina ha ido de la mano junto con el desarrollo y la disponibilidad de los equipos de cómputo que actualmente constituyen una herramienta de fundamental importancia para la implementación de dichos métodos. Como también el desarrollo de los equipos de cómputo ha sido acelerado e ininterrumpido desde 1950, a partir de ese momento y hasta ahora se han ido ampliando las posibilidades de implementación y con ellas las de estudio y aplicación, incluso en aspectos muy básicos de la disciplina como es el caso del presente trabajo.

El estudio de ecuaciones diferenciales mediante técnicas numéricas es una herramienta que en los últimos años ha venido cobrando importancia en el ámbito académico, tanto en el área de matemáticas, como dentro de múltiples

ramas de física e ingeniería a modo de herramienta de investigación. Pero además, actualmente los métodos numéricos ofrecen también aplicaciones en múltiples áreas industriales como son: suministro de energía eléctrica, industrias aeroespacial, automotriz, petroquímica, modelado de explosiones, modelado de inyección de diferentes tipos de materiales licuados como plásticos o aluminio, y fabricación de fibras ópticas, entre muchas otras. Más recientemente pueden encontrarse paquetes de software derivados de este tipo de investigaciones en sitios de internet tales como numeritek.com, century-dynamics.com y muchos otros.

Una de las aplicaciones posibles de este tipo de análisis, y en cierta medida la motivación original del presente trabajo, es el estudio del cambio espectral (corrimiento hacia frecuencias altas) que se presenta en el ruido intenso de banda ancha, al propagarse a lo largo de distancias grandes. Este fenómeno se observa en el ruido que generan las turbinas de un avión jet y su estudio tiene relevancia para no sobrepasar niveles aceptables de ruido ambiental. Sin embargo, el tratamiento de dicho problema requiere, además de una descripción de la propagación del ruido a través de la atmósfera, un estudio de la interacción entre ondas de choque (la llamada turbulencia de Burgers), lo cual rebasa el ámbito del presente trabajo, y es por eso que se discuten únicamente algunos aspectos muy básicos de este tipo de problema.

2. Ondas.

La propagación de ondas es un fenómeno que puede observarse en diferentes áreas de la física. La teoría de ondas estudia estos fenómenos independientemente del sistema físico dentro del cual se produzcan. Precisamente por la diversidad de sistemas físicos que presentan fenómenos ondulatorios es muy difícil precisar lo que es una *onda*, dar una definición de *onda* que abarque todas las posibilidades sin caer en vaguedades. Tal parece que siempre que queremos reconocer alguna propiedad que hemos observado como una característica general de las ondas, la naturaleza nos brinda algún caso que rompe esta regla, y que sin embargo tiene propiedades que también se identifican con movimiento ondulatorio. En el texto “Linear and Nonlinear Waves” de G. B. Whitham [36] el autor discute este problema:

“There appears to be no single precise definition of what exactly constitutes a wave. Various restrictive definitions can be given, but to cover the whole range of wave phenomena it seems preferable to be guided by the intuitive view that a wave is any recognizable signal that is transferred from one part of the medium to another with a recognizable velocity of propagation. The signal may be any feature of the disturbance, such as a maximum or an abrupt change in some quantity, provided that it can be clearly recognizable and its location at any time can be determined.”

“Tal parece que no hay una definición única y precisa de lo que constituye exactamente a una onda. Se pueden dar varias definiciones restrictivas, pero para cubrir todo el rango de fenómenos ondulatorios parece preferible guiarse por la idea intuitiva de que una onda es cualquier señal reconocible que es transferida desde una parte del medio a otra con una velocidad de propagación reconocible. La señal puede ser cualquier aspecto de la perturbación, tal como un máximo o un cambio abrupto en alguna cantidad, siempre que pueda reconocerse claramente y su posición pueda ser determinada en todo momento.”

Además del aspecto principal que destaca Whitham, que es la existencia de una velocidad de propagación, pueden reconocerse en su aseveración otros conceptos necesarios para identificar una onda: un medio, una propiedad del medio, y una perturbación de ésta, lo cual a su vez implica un estado de equilibrio del medio y una perturbación identificable con respecto a este estado de equilibrio (la señal).

Existen dos casos que no serán discutidos en el presente texto y que, sin embargo, vale la pena mencionar con el fin de ampliar la definición de *onda*. El primero se refiere a las ondas electromagnéticas, que se propagan en el vacío, en aparente contradicción a la idea citada de Whitham. En este caso, sin embargo, se trata de ondas electromagnéticas asociadas a las ecuaciones de Maxwell y no de ondas mecánicas, cuya existencia requiere de la presencia de un medio. El vacío juega el papel de un medio que no es un agente propagador de la onda, sino solamente el espacio donde se lleva a cabo el proceso. El segundo caso, quizá aún más enigmático, son las ondas que algunos llaman *ondas de materia* (Feynman [15]), que son soluciones de la ecuación de Schrödinger y describen (bajo una de interpretación) la probabilidad de que una partícula se encuentre en una determinada posición. Aunque no resulta muy clara la relación física que dichas ondas guardan con fenómenos oscilatorios observables, ciertamente su formulación matemática en esta teoría, la mecánica cuántica, tiene notables similitudes con la correspondiente al resto de los fenómenos ondulatorios.

Por otro lado, una de las razones de que la teoría de ondas pueda aplicarse a una diversidad tan grande de sistemas físicos, es que cuando se consideran solamente las perturbaciones de mayor magnitud (de primer orden) a menudo

el problema se reduce a una ecuación modelo bien conocida llamada *ecuación de onda lineal*, en la cual se enfoca el resto de la presente sección.

2.1. Derivación de la ecuación de onda lineal.

Ahora vamos a limitarnos al estudio de las ondas mecánicas en fluidos newtonianos compresibles y sin pérdidas, en reposo, y deduciremos en términos formales una ecuación de onda a partir de las propiedades de estos sistemas. El tipo de medio que queremos describir se puede caracterizar usando tres parámetros definidos como función de la posición en el medio. Estos parámetros son la velocidad \mathbf{u} (velocidad promedio de las moléculas a escala mesoscópica), la presión p y la densidad ρ .

Para deducir la ecuación de onda lineal necesitaremos únicamente tres ecuaciones, que posteriormente vamos a linealizar. Dos de las ecuaciones que necesitaremos corresponden a leyes de conservación:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad \text{Conservación de masa,} \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p + \frac{\eta}{\rho} \nabla^2 \mathbf{u}. \quad \text{Conservación de momento.} \quad (2)$$

La segunda ecuación se conoce como *ecuación de Navier-Stokes* para un fluido incompresible [19], e incluye un término correspondiente a las pérdidas viscosas, donde el coeficiente η se denomina *viscosidad dinámica* y aunque puede depender de ρ y p con frecuencias es considerado constante. Sin embargo como ya hemos señalado, al considerar un medio sin pérdidas despreciaremos dicho término y nos quedaremos con la versión no disipativa que se conoce como *ecuación de Euler*,

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p. \quad (3)$$

La tercera ecuación necesaria es una relación entre ρ y p . Estas dos variables están relacionadas por medio de la ecuación de estado del fluido que se haya escogido y para determinar unívocamente una relación funcional entre ellos es necesario especificar el tipo de proceso. En este caso, siguiendo el texto de Rabinovich [29], consideraremos que se trata de un proceso adiabático y por lo tanto isentrópico. Entonces,

$$\rho = \rho(p).$$

El siguiente paso es expresar las variables como pequeñas perturbaciones a partir de sus correspondientes valores en estado de equilibrio. Cuando el sistema está en equilibrio sus variables tienen los siguientes valores:

$$\begin{aligned} p &= p_0, \\ \rho &= \rho_0 = \rho(p_0), \\ \mathbf{u} &= \mathbf{u}_0 = \mathbf{0} \quad (\text{El fluido está originalmente en reposo}). \end{aligned}$$

Podemos ahora escribir la relación que guardan las pequeñas perturbaciones $p_1, \rho_1, \mathbf{u}_1$ con respecto a los valores de

equilibrio del siguiente modo:

$$\begin{aligned} p(\mathbf{x}, t) &= p_0 + p_1(\mathbf{x}, t), \\ \rho(\mathbf{x}, t) &= \rho_0 + \rho_1(\mathbf{x}, t), \\ \mathbf{u}(\mathbf{x}, t) &= \mathbf{u}_1(\mathbf{x}, t). \end{aligned} \quad (4)$$

La idea de definir estas nuevas variables es aprovechar el hecho de que son pequeñas con respecto a los valores de referencia

$$\begin{aligned} |\rho_1| &\ll \rho_0, \\ |p_1| &\ll p_0, \\ |\mathbf{u}_1| &\ll 1. \end{aligned} \quad (5)$$

Linealizando los términos que aparecen en (1) y (3), es decir, sustituyendo (4) y eliminando factores de orden 2 o más en las variables arriba descritas (pues son de orden de magnitud más pequeños) se obtiene:

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{\partial \rho_1}{\partial t}, \\ \nabla \cdot (\rho \mathbf{u}) &\simeq \nabla \cdot (\rho_0 \mathbf{u}) = \rho_0 \nabla \cdot \mathbf{u}, \\ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} &\simeq \frac{\partial \mathbf{u}}{\partial t}, \\ \frac{1}{\rho} &\simeq \frac{1}{\rho_0}. \end{aligned}$$

Sustituyendo de nuevo, ahora en las ecuaciones de conservación (1) y (3) tenemos

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{u} = 0, \quad (6)$$

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{\nabla p_1}{\rho_0}, \quad (7)$$

que son las formas linealizadas de las ecuaciones de conservación de masa y momento lineal, respectivamente.

Derivando la ecuación (6) parcialmente con respecto a t , siempre y cuando $\mathbf{u} \in C^2$ de modo que se pueda intercambiar el orden de derivación, se tiene

$$\frac{\partial^2 \rho_1}{\partial t^2} + \rho_0 \nabla \cdot \frac{\partial \mathbf{u}}{\partial t} = 0,$$

por lo que sustituyendo ahora en esta ecuación la ecuación (7) para eliminar el factor que contiene a \mathbf{u} , se obtiene

$$\begin{aligned} \frac{\partial^2 \rho_1}{\partial t^2} + \rho_0 \nabla \cdot \left(-\frac{\nabla p_1}{\rho_0} \right) &= 0, \\ \frac{\partial^2 \rho_1}{\partial t^2} - \nabla^2 p_1 &= 0. \end{aligned} \quad (8)$$

Finalmente recurrimos a la ecuación de estado isentrópica que relaciona ρ y p , $\rho = \rho(p)$. Esta relación se puede

expresar por medio de una serie de Taylor de primer orden en torno a las variables antes definidas,

$$\rho_0 + \rho_1 \simeq \rho(p_0) + \left(\frac{d\rho}{dp}\right)_S \cdot (p_1) ,$$

donde se entiende que la derivada de ρ con respecto a p está evaluada precisamente en p_0 . Además, $\rho(p_0)$ es precisamente ρ_0 , por lo que

$$\rho_1 = \left(\frac{d\rho}{dp}\right)_S \cdot p_1 . \quad (9)$$

Si ahora sustituimos este resultado en la ecuación (8), se obtiene la ecuación de onda lineal,

$$\frac{1}{c^2} \frac{\partial^2 p_1}{\partial t^2} - \nabla^2 p_1 = 0 ,$$

donde la constante c está definida por

$$c = \left(\frac{d\rho}{dp}\right)_S^{-1/2} ,$$

y se identificará más tarde con la velocidad de propagación de la perturbación de presión.

Un aspecto que se puede observar en este momento es que la relación entre ρ y p (determinada por las propiedades del medio) está directamente relacionada con la velocidad de propagación, de modo que se puede hacer una comparación entre el valor que resulte de medir c experimentalmente a partir de cantidades termodinámicas, y el valor que se obtenga midiendo la velocidad de propagación.

Para simplificar la notación a partir de este punto, a menos de que se indique lo contrario, las variables ρ , p y u siempre representan perturbaciones respecto al estado de equilibrio y se escriben sin subíndices. Entonces, en el caso de la presión la ecuación de onda lineal es

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0 . \quad (10)$$

Es posible mostrar que otras magnitudes que describen al fluido obedecen la misma ecuación. De entrada, bajo la suposición que se ha planteado en la ecuación (9) de que ρ y p guardan una relación lineal, puede observarse directamente sustituyendo en la ecuación (10) que la densidad efectivamente satisface la ecuación de onda:

$$\nabla^2 \rho - \frac{1}{c^2} \frac{\partial^2 \rho}{\partial t^2} = 0 .$$

Para el caso de \mathbf{u} observamos que se trata de un vector y que la ecuación de onda se ha obtenido previamente como una ecuación escalar; sin embargo, podemos mostrar que las componentes de \mathbf{u} , que son escalares, satisfacen la ecuación de onda. Primero sustituimos la ecuación (9) en la ecuación (7) y tenemos

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{c^2 \nabla \rho}{\rho_0} .$$

Posteriormente, derivamos con respecto a t y reordenamos los términos para encontrar

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} \left(-\frac{\rho_0}{c^2}\right) = \nabla \frac{\partial \rho}{\partial t} ; \quad (11)$$

por otro lado tomamos el gradiente en ambos lados de la ecuación (6),

$$\nabla \frac{\partial \rho}{\partial t} + \rho_0 \nabla (\nabla \cdot \mathbf{u}) = 0 ,$$

y sustituimos la ecuación (11) en el resultado anterior para obtener

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} \left(-\frac{\rho_0}{c^2} \right) + \rho_0 \nabla (\nabla \cdot \mathbf{u}) = 0 .$$

Utilizando una conocida identidad, $\nabla (\nabla \cdot \mathbf{A}) = \nabla^2 \mathbf{A} + \nabla \times (\nabla \times \mathbf{A})$, se obtiene

$$\frac{-1}{c^2} \frac{\partial^2 \mathbf{u}}{\partial t^2} + \nabla^2 \mathbf{u} + \nabla \times (\nabla \times \mathbf{u}) = 0 .$$

Tomando en cuenta que el término $\nabla^2 \mathbf{u}$ es precisamente un vector cuyas componentes son los laplacianos de las componentes de \mathbf{u} ,

$$\nabla^2 \mathbf{u} = (\nabla^2 u_x, \nabla^2 u_y, \nabla^2 u_z) ,$$

prácticamente hemos escrito la ecuación de onda para cada una de las componentes de \mathbf{u} , pero queda un término que involucra al rotacional de \mathbf{u} . Eliminaremos este término bajo la argumentación de que nuestro flujo es irrotacional, y no presenta vórtices. Para un flujo irrotacional, $\nabla \times \mathbf{u} = 0$, y entonces

$$\frac{-1}{c^2} \frac{\partial^2 \mathbf{u}}{\partial t^2} + \nabla^2 \mathbf{u} = 0 ,$$

o, lo que es lo mismo, cada una de las componentes de \mathbf{u} satisface la ecuación de onda:

$$\begin{aligned} \nabla^2 u_x - \frac{1}{c^2} \frac{\partial^2 u_x}{\partial t^2} &= 0 , \\ \nabla^2 u_y - \frac{1}{c^2} \frac{\partial^2 u_y}{\partial t^2} &= 0 , \\ \nabla^2 u_z - \frac{1}{c^2} \frac{\partial^2 u_z}{\partial t^2} &= 0 . \end{aligned}$$

Para los desarrollos posteriores, a menos de que se especifique lo contrario, trabajaremos en una sola dimensión y en este caso la magnitud de \mathbf{u} es igual a su única componente, u , por lo que satisface la ecuación de onda

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 . \quad (12)$$

Sin embargo, es importante notar que no hemos demostrado que la magnitud del vector \mathbf{u} satisfaga la ecuación de onda para el caso tridimensional, y de hecho no la satisface.

Para ciertos desarrollos resulta más cómodo tener las ecuaciones en términos de u y como en la mayoría de los casos aquí tratados será suficiente trabajar en una dimensión espacial podremos usar esta variable u sin problemas; cuando por alguna razón p o ρ resulten más convenientes, así se hará notar.

2.2. Caso unidimensional: ecuación de advección y solución de D'Alembert.

Con el fin de describir algunas propiedades de la solución general de la ecuación de onda expondremos un desarrollo por medio del cual se deduce la solución del caso unidimensional en un dominio infinito. La ecuación de onda en este caso (12) se expresa de este modo:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 .$$

Se hace un cambio de variables:

$$\begin{aligned} \eta &= x - ct , \\ \xi &= x + ct , \end{aligned}$$

por lo que

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= c^2 \left(\frac{\partial^2 u}{\partial \xi^2} - 2 \frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \eta^2} \right) , \\ \frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial \xi^2} + 2 \frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \eta^2} . \end{aligned}$$

Al sustituir estas expresiones en la ecuación (12) tenemos

$$\frac{\partial^2 u}{\partial \eta \partial \xi} = 0$$

que se puede resolver por integración directa. Primero integramos con respecto a ξ y tenemos

$$\frac{\partial u}{\partial \eta} = F_1(\eta) ,$$

donde F_1 es cualquier función de η . Luego integramos con respecto a η para tener

$$u = \int^{\eta} F_1(z) dz + f_2(\xi) ,$$

donde f_2 también es una función arbitraria de ξ . Sin pérdida de generalidad y sólo con el fin de aligerar la notación definimos otra función f_1 mediante

$$f_1(\eta) = \int^{\eta} F_1(z) dz ,$$

y entonces la solución general es

$$u = f_1(\eta) + f_2(\xi) ,$$

que, expresada en las variables originales tiene la siguiente forma

$$u(x, t) = f_1(x - ct) + f_2(x + ct) .$$

Para interpretar cualitativamente esta solución supongamos que la función f_2 es idénticamente cero, de modo

que

$$u(x, t) = f_1(x - ct) .$$

Para $t = 0$ la solución u es precisamente f_1 :

$$u(x, 0) = f_1(x) .$$

Para cualquier tiempo posterior t la solución u será la misma función f_1 pero desplazada hacia la derecha un intervalo igual a ct , de modo que decimos que u describe una onda que viaja sin deformarse hacia la derecha con una velocidad constante c . Del mismo modo, en el caso de que f_1 sea nula, u describe una onda que viaja hacia la izquierda sin deformarse con una velocidad de la misma magnitud c . El caso general lo podemos entender como la suma de dos ondas que viajan en direcciones opuestas con la misma rapidez c .

2.2.1. Otra manera de expresar la ecuación de onda: ecuación de advección.

Otra manera de observar este carácter de ondas que viajan en direcciones opuestas se logra al escribir la ecuación de onda (12) en las siguientes formas:

$$\begin{aligned} 0 &= \frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} \\ &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} + \frac{1}{c} \frac{\partial u}{\partial t} \right) - \frac{1}{c} \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} + \frac{1}{c} \frac{\partial u}{\partial t} \right) , \end{aligned} \quad (13)$$

$$= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} - \frac{1}{c} \frac{\partial u}{\partial t} \right) + \frac{1}{c} \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} - \frac{1}{c} \frac{\partial u}{\partial t} \right) . \quad (14)$$

Estas formas son equivalentes siempre que la función u sea suficientemente suave (de clase C^2) como para que $\frac{\partial^2 u}{\partial x \partial t} = \frac{\partial^2 u}{\partial t \partial x}$. Mediante esta expresión es fácil notar que si una función es solución de alguna de las siguientes ecuaciones

$$\frac{\partial u}{\partial x} + \frac{1}{c} \frac{\partial u}{\partial t} = 0 , \quad (15)$$

$$\frac{\partial u}{\partial x} - \frac{1}{c} \frac{\partial u}{\partial t} = 0 , \quad (16)$$

entonces también será solución de la ecuación de onda unidimensional. Además, es posible demostrar que si se tiene una función $u_1(x, t)$ solución de la ecuación (15), y una función $u_2(x, t)$ solución de la ecuación (16), la suma de ambas

$$u = u_1 + u_2 ,$$

también será solución de la ecuación de onda unidimensional. Desde luego las soluciones generales de las ecuaciones (15) y (16) son precisamente las ondas que viajan en ambas direcciones, es decir,

$$u_1(x, t) = f_1(x - ct) ,$$

$$u_2(x, t) = f_2(x + ct) ,$$

donde f_1 y f_2 son funciones de \mathbb{R} en \mathbb{R} .

La importancia de reconocer que la ecuación de onda unidimensional se puede expresar usando las ecuaciones

(15) y (16) radica en que mediante la solución general de cualquiera de estas últimas (que es más fácil de encontrar), por simetría, se puede encontrar la solución general de la otra y luego a su vez la solución general de la ecuación de onda. Las ecuaciones (15) y (16) son la misma ecuación salvo un signo que se refiere a dirección de la propagación; a esta ecuación se le conoce como *ecuación de advección* y sin pérdida de generalidad usaremos en adelante la versión cuyas soluciones, cuando $c > 0$, se propagan hacia la derecha

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0. \quad (17)$$

La idea detrás de este procedimiento es la siguiente: la ecuación de onda (lineal o no-lineal) debe describir la propagación de una onda en cualquier dirección, pero como el medio que estamos modelando es isotrópico las características de la propagación no deben depender de la dirección de propagación, por lo que podemos usar una ecuación que describa la propagación en una sola dirección, y después la generalización de los resultados no debe presentar problemas.

Una manera abreviada de escribir las expresiones (13) y (14) es

$$0 = \left(\frac{\partial}{\partial x} - \frac{1}{c} \frac{\partial}{\partial t} \right) \left(\frac{\partial}{\partial x} + \frac{1}{c} \frac{\partial}{\partial t} \right) u, \quad (18)$$

$$= \left(\frac{\partial}{\partial x} + \frac{1}{c} \frac{\partial}{\partial t} \right) \left(\frac{\partial}{\partial x} - \frac{1}{c} \frac{\partial}{\partial t} \right) u, \quad (19)$$

donde es importante hacer notar que los paréntesis **no** denotan producto sino composición. En caso de que se quisiera escribir un producto de este tipo, la ecuación correspondiente sería

$$\begin{aligned} 0 &= \left(\frac{\partial u}{\partial x} - \frac{1}{c} \frac{\partial u}{\partial t} \right) \left(\frac{\partial u}{\partial x} + \frac{1}{c} \frac{\partial u}{\partial t} \right) \\ &= \left(\frac{\partial u}{\partial x} \right)^2 - \frac{1}{c^2} \left(\frac{\partial u}{\partial t} \right)^2, \end{aligned} \quad (20)$$

cuyas soluciones ciertamente son soluciones de la ecuación (15) o de la ecuación (16) pero si tenemos una función que es la suma de una solución de cada una de estas dos ecuaciones, esta función no necesariamente será solución de la ecuación (20). Es decir, si

$$u = f_1(x - ct) + f_2(x + ct),$$

entonces

$$\left(\frac{\partial u}{\partial x} - \frac{1}{c} \frac{\partial u}{\partial t} \right) \left(\frac{\partial u}{\partial x} + \frac{1}{c} \frac{\partial u}{\partial t} \right) = 4f_1'(x - ct)f_2'(x + ct),$$

lo cual no necesariamente es cero.

2.2.2. Condiciones iniciales.

El siguiente paso es precisar cuáles funciones f_1 y f_2 deben usarse para describir una situación específica. Como existen infinitas maneras de lograr una condición inicial de u mediante la suma de dos funciones arbitrarias, además de especificar $u(x, 0)$ será necesario especificar, para el mismo $t = 0$, el valor de su derivada parcial con respecto al

tiempo. Supongamos entonces que se dan, inicialmente,

$$\begin{aligned} u(x, 0) &= g(x) , \\ \left. \frac{\partial u}{\partial t} \right|_{t=0} &= G(x) . \end{aligned}$$

Buscaremos entonces expresar la solución general $u(x, t)$ en términos de g y G .

Por un lado debemos tener, debido a la primera condición,

$$f_1(x) + f_2(x) = g(x) \quad (21)$$

y por otro lado, debido a la segunda condición,

$$c(-f_1'(x) + f_2'(x)) = G(x) . \quad (22)$$

Por lo tanto, integrando la ecuación (22) con respecto a una variable muda z desde un punto arbitrario x_0 hasta x tenemos

$$(-f_1(x) + f_1(x_0)) + (f_2(x) - f_2(x_0)) = \frac{1}{c} \int_{x_0}^x G(z) dz .$$

Entonces se tiene

$$f_1(x) = -\frac{1}{c} \int_{x_0}^x G(z) dz + f_1(x_0) + f_2(x) - f_2(x_0) .$$

Sustituyendo la ecuación (21) en el resultado anterior

$$\begin{aligned} f_1(x) &= -\frac{1}{c} \int_{x_0}^x G(z) dz + f_1(x_0) + g(x) - f_1(x) - f_2(x_0) \\ &= -\frac{1}{2c} \int_{x_0}^x G(z) dz + \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x) - \frac{1}{2}f_2(x_0) , \end{aligned}$$

y usando nuevamente la ecuación (21),

$$\begin{aligned} f_2(x) &= g(x) + \frac{1}{2c} \int_{x_0}^x G(z) dz - \frac{1}{2}f_1(x_0) - \frac{1}{2}g(x) + \frac{1}{2}f_2(x_0) \\ &= \frac{1}{2c} \int_{x_0}^x G(z) dz - \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x) + \frac{1}{2}f_2(x_0) . \end{aligned}$$

Ahora evaluamos en los argumentos que les corresponden dentro de la solución general,

$$\begin{aligned} f_1(x - ct) &= -\frac{1}{2c} \int_{x_0}^{x-ct} G(z) dz + \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x - ct) - \frac{1}{2}f_2(x_0) , \\ f_2(x + ct) &= \frac{1}{2c} \int_{x_0}^{x+ct} G(z) dz - \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x + ct) + \frac{1}{2}f_2(x_0) , \end{aligned}$$

y finalmente podemos expresar u en términos de sus condiciones iniciales

$$\begin{aligned} u(x, t) &= \left(-\frac{1}{2c} \int_{x_0}^{x-ct} G(z) dz + \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x-ct) - \frac{1}{2}f_2(x_0) \right) \\ &\quad + \left(\frac{1}{2c} \int_{x_0}^{x+ct} G(z) dz - \frac{1}{2}f_1(x_0) + \frac{1}{2}g(x+ct) + \frac{1}{2}f_2(x_0) \right) \\ &= \frac{1}{2c} \int_{x-ct}^{x+ct} G(z) dz + \frac{1}{2}g(x-ct) + \frac{1}{2}g(x+ct) . \end{aligned} \quad (23)$$

Esta última expresión se conoce como solución de D'Alembert [5].

2.2.3. Curvas características.

El *método de características*, formulado por Lagrange, es un método ampliamente usado para encontrar soluciones a ecuaciones diferenciales parciales. Para una ecuación cuasilineal de primer orden, como las que trabajaremos en el presente texto, escrita en su forma más general,

$$a(x, t, u) u_t + b(x, t, u) u_x = 0 , \quad (24)$$

con una condición inicial dada,

$$u(x, 0) = g(x) , \quad (25)$$

este método consiste en encontrar aquellas trayectorias, $x(z), t(z)$, sobre las cuales se cumple

$$\frac{dt}{dz} = a(x, t, u) , \quad (26)$$

$$\frac{dx}{dz} = b(x, t, u) . \quad (27)$$

A estas trayectorias se les conoce con el nombre de *curvas características* o simplemente *características*. Una descripción más general puede encontrarse por ejemplo en el texto de Debnath [12].

La ventaja que ofrecen las curvas características puede observarse al tomar la derivada de una función u sobre alguna de ellas mismas:

$$\frac{du}{dz} = u_t \frac{dt}{dz} + u_x \frac{dx}{dz} .$$

Si ahora se sustituyen en esta expresión las ecuaciones (26), (27) y (24) observamos que

$$\frac{du}{dz} = 0 .$$

En otras palabras, para estos casos u es una constante sobre las características, de modo que la solución de (24) debería poder encontrarse resolviendo las ecuaciones (26) y (27) para encontrar las curvas características, y luego determinando el valor de u que corresponde a cada una de éstas por medio de su intersección con la condición inicial $g(x)$. Es decir, si la característica interseca al eje x en el punto $x = x_0, t = 0$, el valor de u que le corresponde debe ser precisamente $g(x_0)$.

En los casos estudiados en el presente texto, nos será posible parametrizar con respecto a la misma t de modo

que las características estarán descritas mediante

$$x = X(t) ,$$

sobre las cuales

$$u(X(t), t) = g(x_0) \quad \text{para toda } t .$$

Cuando tenemos una ecuación diferencial parcial del tipo de la ecuación de onda lineal, que expresa la propagación de una perturbación con todas sus propiedades, pero en una sola dirección, como la que se ha obtenido en la sección 2.2.1, las curvas características son una buena manera de visualizar estas propiedades: si las características se aproximan unas a otras tenemos pulsos que se hacen estrechos, y por el contrario, si se alejan entre si tenemos pulsos que se ensanchan. Un aumento en la pendiente $X'(t)$ corresponde a un aumento en la velocidad de propagación y viceversa. Una pendiente positiva corresponde a una propagación hacia la derecha y una pendiente negativa corresponde al sentido contrario. Un poco menos intuitivos son los casos donde las características se intersectan, dichos casos dan lugar a perfiles multivaluados y requieren un tratamiento especial, que será abordado en la sección 3.

En este momento sólo analizaremos las características que corresponden a la ecuación de advección (17)

$$u_t + cu_x = 0 .$$

Escribiendo las ecuaciones (26) y (27) en este caso obtenemos

$$\frac{dx}{dz} = c , \tag{28}$$

$$\frac{dt}{dz} = 1 , \tag{29}$$

por lo que las características en este caso, parametrizadas con respecto a t , tienen la forma

$$x = X(t) = ct + x_0 .$$

Ya que para esta ecuación diferencial se ha calculado previamente (en la sección 2.2.1) la solución analítica $u(x, t) = f(x - ct)$, puede verificarse la consistencia de estos dos resultados observando que sobre las características obtenidas u efectivamente sea constante,

$$\begin{aligned} u(X(t), t) &= f(X(t) - ct) \\ &= f(x_0 + ct - ct) \\ &= f(x_0) . \end{aligned}$$

Hemos encontrado pues que las curvas características en este caso son rectas de pendiente c positiva, tal y como se esperaba. Estas mismas se muestran en la figura 1a. En la figura 1b se ha añadido un eje correspondiente a la variable dependiente u y se puede observar como se propaga una condición inicial sin deformarse.

La importancia del análisis por características será evidente cuando se analicen casos donde las ondas se deforman mientras viajan, ya que debido a la deformación misma no será posible expresar las soluciones mediante una

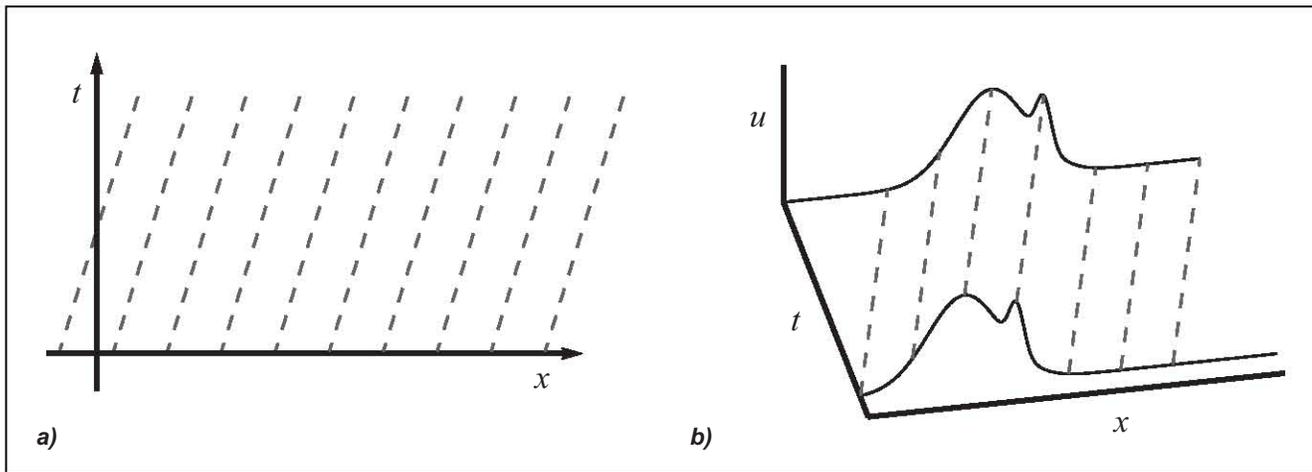


Figura 1: Propagación lineal de una onda, las líneas punteadas en ambos casos son las curvas *características*. El inciso **a** es una vista superior del inciso **b**.

función de la forma $f(x \pm ct)$ para todo tiempo. No será entonces posible describir la propagación de la condición inicial como una función que simplemente se desplaza sin deformarse; en lugar de eso describimos la propagación de cada uno de los puntos de la condición inicial por medio de una curva característica.

2.3. Principio de superposición.

El hecho de que la ecuación de onda (10)

$$\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0$$

sea lineal, formalmente quiere decir que dadas $u_1(x, t)$ y $u_2(x, t)$ dos soluciones a la ecuación, para cualesquiera a_1 y a_2 escalares, la función

$$u = a_1 u_1 + a_2 u_2,$$

que consiste sencillamente de una combinación lineal de dos soluciones, también es una solución. La demostración de dicha propiedad es sencilla:

$$\begin{aligned} \nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} &= \nabla^2 (a_1 u_1 + a_2 u_2) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} (a_1 u_1 + a_2 u_2) \\ &= a_1 \nabla^2 u_1 + a_2 \nabla^2 u_2 - \left(a_1 \frac{1}{c^2} \frac{\partial^2 u_1}{\partial t^2} + a_2 \frac{1}{c^2} \frac{\partial^2 u_2}{\partial t^2} \right) \\ &= a_1 \left(\nabla^2 u_1 - \frac{1}{c^2} \frac{\partial^2 u_1}{\partial t^2} \right) + a_2 \left(\nabla^2 u_2 - \frac{1}{c^2} \frac{\partial^2 u_2}{\partial t^2} \right) \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Por un lado esta propiedad nos dice que las propiedades de la propagación no dependen de la amplitud de la onda. En otras palabras, una vez que hemos encontrado una solución podemos amplificarla positiva o negativamente multiplicándola por cualquier escalar y el resultado será igualmente una solución de la ecuación. Por otra parte nos

dice que las interacciones entre las ondas son lineales; es decir, el proceso de propagación es independiente de que haya otro proceso de propagación ocurriendo en el mismo tiempo en el mismo lugar.

Existe también una relación entre este principio de superposición y el hecho de que todas las ondas se propaguen a la misma velocidad. Primero hagamos una observación: si la condición inicial de una onda que viaja sin deformarse a una determinada velocidad se puede expresar como la suma de dos funciones, entonces las perturbaciones correspondientes a estas funciones deben viajar a la misma velocidad que la primera. Ahora, si suponemos que existen dos perturbaciones que viajan sin deformarse a velocidades diferentes encontramos rápidamente una contradicción: al expresar la condición inicial de la primera como la suma de la condición inicial de la segunda y otra función (la diferencia entre ambas) y apoyados en el hecho de que los procesos de propagación son independientes, habría una parte de la perturbación viajando a la velocidad correspondiente a la segunda, y no a la velocidad que se había supuesto en un principio.

El principio de superposición nos permite analizar la solución general de la ecuación en el sentido literal de la palabra, entenderla como un número finito de partes independientes que se estudian por separado y luego encontrar los resultados generales sumando los resultados de todas las partes.

2.4. Dispersión y difusión.

Cuando se observa que en un mismo medio *diferentes* ondas se propagan con diferentes velocidades de fase, se dice que el medio es *dispersivo*. Una definición formal del concepto de dispersión es difícil porque depende del concepto de onda que, como se ha visto al principio de esta sección 2, presenta problemas. Whitham [36] menciona una definición que resulta útil y consiste en llamar *dispersivos* a los casos donde la ecuación diferencial

$$F[u] = 0$$

admite soluciones $u(x, t)$ de la forma

$$u(x, t) = a \cos(\kappa x - \omega t + \phi), \quad (30)$$

que cumplen con las siguientes condiciones:

- ω es una función real de κ ; esto es $\omega = \omega(\kappa)$ (esta función se conoce como *relación de dispersión*),
- el valor de la velocidad de fase $c = \frac{\omega}{\kappa}$ no es independiente del valor de κ .

Luego se puede extender la definición a los casos que han sido excluidos siempre y cuando *cuantitativamente* sus respectivas propagaciones sean análogas a los casos que caben dentro de la definición mencionada. Para el caso de la ecuación de advección es posible aplicar la definición estricta sugerida por Whitham, dado que se satisfacen las condiciones por ella impuestas.

Por otro lado, desde el punto de vista físico el término *difusión*, o *disipación* [21], se refiere a los procesos en los cuales la energía tiende a distribuirse uniformemente en todo el sistema. Entonces en los sistemas *difusivos* las perturbaciones al propagarse pierden amplitud gradualmente hasta desaparecer, debido a que dichas perturbaciones llevan consigo una cierta cantidad finita de energía. Por lo tanto, las soluciones de una ecuación diferencial que modele un sistema difusivo deben, a medida que aumenta el tiempo, tender a los estados de equilibrio del sistema. Típicamente estas soluciones involucran decaimiento exponencial porque comúnmente, bajo alguna manipulación,

dichas ecuaciones pueden reducirse, al menos en parte, a la *ecuación de difusión*

$$u_t = Ku_{xx} ,$$

cuyas soluciones son de la forma

$$u(x, t) = e^{-\beta t} a \cos(\kappa x + \phi) , \quad (31)$$

y decrecen exponencialmente con respecto a la variable temporal.

Es notable que las soluciones expresadas en las ecuaciones (30) y (31) pueden ser entendidas como la parte real de una exponencial compleja

$$ae^{i(\kappa x - \omega t + \phi)} ,$$

donde a la forma expresada en la ecuación (30) corresponde una relación de dispersión $\omega(\kappa)$ puramente real y para la forma expresada en la ecuación (31) la función $\omega(\kappa)$ debe ser puramente imaginaria, en este último caso se tendría $\omega(\kappa) = i\beta(\kappa)$. Esto nos lleva a notar que bien puede suceder que $\omega(\kappa)$ tenga, tanto parte real como parte imaginaria distintas de cero, y eso significa que pueden existir propagaciones que manifiesten difusión y dispersión al mismo tiempo.

En ocasiones es difícil distinguir entre dispersión y difusión porque el perfil de una onda al propagarse puede ser deformado y cambiar su amplitud por ambas razones. Por eso es importante precisar la definición que se usa al hablar de dispersión, por ejemplo la definición rigurosa de Whitham, que está basada en soluciones del tipo senoidal, permite, siempre y cuando se cumpla el principio de superposición en el medio, hacer una descomposición del perfil, en este caso por medio de una serie de Fourier, y determinar si hay difusión presente observando si decrece la amplitud de las componentes senoidales.

3. Ondas de choque.

Desde el punto de vista de la física las *ondas de choque* son regiones estrechas que se propagan, en las cuales cambia abruptamente (pero de modo continuo) alguna (o algunas) de las propiedades físicas del medio. Estas regiones pueden ser tan estrechas como el orden de magnitud del *camino libre medio* de las moléculas, esto es 10^{-7} m en la tierra; un tratamiento detallado puede encontrarse en los textos de Toro [33], Whitham [36] y Landau y Lifshitz [19]. Comúnmente la primera aproximación que se usa para modelar matemáticamente dichas ondas de choque es una descripción por medio de funciones discontinuas; es así como lo haremos en el presente texto.

3.1. Ecuación de Burgers no viscosa.

Tal vez el ejemplo más sencillo de ecuaciones diferenciales parciales cuyas soluciones presentan discontinuidades correspondientes a ondas de choque es la ecuación de Burgers, en particular su versión no viscosa. Una manera de obtener esta ecuación es planteando una ley de conservación, por ejemplo, conservación de masa. Supóngase que en un determinado medio unidimensional la densidad de masa es ρ , denotando con q su correspondiente flujo; entonces tanto ρ como q pueden ser consideradas como funciones de las variables espacial x y temporal t ,

$$\begin{aligned}\rho &= \rho(x, t), \\ q &= q(x, t).\end{aligned}$$

Estas variables se relacionan entre sí por medio de

$$\frac{d}{dt} \int_{x_1}^{x_2} \rho dx = q(x_2, t) - q(x_1, t), \quad (32)$$

lo que significa que en cualquier intervalo delimitado por dos puntos x_1, x_2 , el cambio en la masa contenida en dicho intervalo es igual a la diferencia entre los flujos de masa en los extremos del mismo intervalo, donde se han tomado como positivos los flujos de izquierda a derecha. A la expresión de (32) se le conoce como ley de conservación de masa en forma integral. Si ahora hacemos tender $x_1 \rightarrow x_2$ obtenemos su forma diferencial

$$\rho_t + q_x = 0. \quad (33)$$

Como ésta es una sola ecuación diferencial que involucra las dos variables ρ y q , se necesita otra relación entre estos mismos parámetros para determinar el problema y poder buscar una solución. Una solución de este problema, primero planteada por Whitham [36], es considerar q como función de ρ , $q = q(\rho)$, lo cual implica que

$$q_x = q'(\rho) \rho_x, \quad (34)$$

y por lo tanto, al sustituir en (33),

$$\rho_t + q'(\rho) \rho_x = 0, \quad (35)$$

donde, si bien siguen apareciendo las dos variables mencionadas, podemos evitar las apariciones de q aproximándola por polinomios en ρ (lo cual equivale a tomar la serie de Taylor truncada de una función $q(\rho)$ general). Si q es un

polinomio de primer orden de ρ se tiene

$$q(\rho) = A + B\rho ,$$

donde A y B son constantes, entonces $q'(\rho) = B$ y la ecuación (35) adquiere la forma $\rho_t + B\rho_x = 0$, que es precisamente la ecuación de advección que se ha descrito en la sección 2.2.1.

Si vamos un poco más adelante y usamos un polinomio de segundo orden tenemos

$$\begin{aligned} q(\rho) &= A + B\rho + C\rho^2 , \\ q'(\rho) &= B + 2C\rho , \end{aligned} \tag{36}$$

con lo que la ecuación (35) queda

$$\rho_t + (B + 2C\rho)\rho_x = 0 . \tag{37}$$

Para simplificar esta ecuación definimos una nueva variable φ linealmente relacionada con ρ :

$$\varphi = B + 2C\rho . \tag{38}$$

Luego, como esto implica que las derivadas parciales de φ son:

$$\begin{aligned} \varphi_x &= 2C\rho_x , \\ \varphi_t &= 2C\rho_t , \end{aligned}$$

entonces la nueva variable φ debe satisfacer

$$\varphi_t + \varphi\varphi_x = 0 . \tag{39}$$

A esta ecuación se le conoce como *ecuación de Burgers no viscosa*. Procediendo ahora en sentido opuesto, es posible observar que la variable φ satisface a su vez una ecuación del tipo (32), correspondiente a la conservación de φ en su forma integral. Sustituyendo (36) y (38) en (32), después de un poco de álgebra se obtiene

$$\begin{aligned} \frac{d}{dt} \int_{x_1}^{x_2} \varphi dx &= \frac{1}{2} (\varphi(x_2, t))^2 - \frac{1}{2} (\varphi(x_1, t))^2 , \\ &= \tilde{q}(x_2, t) - \tilde{q}(x_1, t) , \end{aligned} \tag{40}$$

donde se observa que el flujo \tilde{q} correspondiente a la variable φ puede expresarse como $\tilde{q}(x, t) = \frac{1}{2} (\varphi(x, t))^2$.

Entonces, la ecuación de Burgers no viscosa puede obtenerse al plantear una ley de conservación, de una determinada magnitud en un medio, aproximando el flujo a segundo orden como una función de ρ . Una ecuación de Burgers unidimensional más completa que incluya un término viscoso se escribe

$$\varphi_t + \varphi\varphi_x = \nu\varphi_{xx} ,$$

y puede obtenerse considerando q como una función no solamente de ρ sino también de ρ_x , [36].

Para obtener la solución analítica de la ecuación de Burgers no viscosa (39) sujeta a una condición inicial $\varphi(x, 0) = g(x)$, se procede por medio de curvas características, que han sido brevemente descritas en la sección

2.2.3. Buscamos trayectorias $x(z)$, $t(z)$ sobre las cuales

$$\frac{dx}{dz} = \varphi, \quad (41)$$

$$\frac{dt}{dz} = 1, \quad (42)$$

de modo que las características son rectas de pendiente φ que podemos parametrizar con respecto a la variable t como

$$X(t) = \varphi t + x_0,$$

donde x_0 es el valor de x correspondiente a la intersección de la característica con el eje x . Sobre estas curvas, como se hecho notar en la sección 2.2.3, el valor de φ es una constante,

$$\varphi(X(t), t) = \varphi(x_0, 0) = g(x_0) \quad \text{para toda } t, \quad (43)$$

de modo que podemos reescribirlas como

$$X(t) = g(x_0) t + x_0. \quad (44)$$

La diferencia del caso de la ecuación de Burgers no viscosa con respecto al de la ecuación de advección es que la pendiente de las características ahora depende del valor de la condición inicial $g(x)$ en el punto donde se intersecta con la característica, por lo que, salvo en el caso de una condición inicial constante, las características serán no paralelas y esto significa que existirán puntos por donde pudiera pasar más de una característica y puntos por donde no pase ninguna de ellas. Como será mostrado en la siguiente sección, el primero de estos casos (zonas donde las características se intersectan) da lugar a soluciones de tipo discontinuo.

La exposición anterior con respecto a la ecuación de Burgers no viscosa para ρ pretende mostrar algo de la naturaleza de la ecuación y el tipo de situaciones donde puede surgir. Sin embargo, en las secciones posteriores consideraremos dicha ecuación en la variable u por lo cual es necesario hacer las siguientes observaciones.

Como la relación que guardan el flujo, la densidad y la velocidad es

$$q = \rho u,$$

entonces al sustituir dicha expresión en la ecuación (36) tenemos

$$\rho u = A + B\rho + C\rho^2 \quad (45)$$

ahora, con el fin de obtener una ecuación diferencial para la variable u debería despejarse ρ de la ecuación (45) y luego sustituir el resultado en la expresión (37), si al realizar dicho procedimiento se consideran las tres constantes A , B y C distintas de cero, se obtiene una ecuación diferencial relativamente complicada, sin embargo, si por el contrario se asume que $A = 0$ el resultado es

$$u_t + (2u - B)u_x = 0,$$

que como se ha visto antes puede reducirse a la ecuación de Burgers no viscosa por medio de un cambio de variable

sencillo, que en este caso sería

$$\varphi = 2u - B.$$

Dicho cambio de variable equivale simplemente a medir la velocidad de desde un sistema de referencia distinto (y además con un cambio de escala), es decir, en este caso φ es una representación de la velocidad promedio de las partículas a escala mesoscópica, tan buena como u .

Entonces una manera de obtener la ecuación de Burgers para la variable u consiste en pedir que la variable ρ cumpla una ley de conservación, como en la ecuación (32), e imponer dos restricciones: que el flujo q guarde una relación cuadrática con ρ , y que dicha relación carezca de término a orden cero, $A = 0$, lo cual puede ser interpretado como limitarse a los casos donde el flujo se anula cuando la densidad se aproxima a cero.

3.2. Condiciones de Rankine-Hugoniot.

Como se ha mencionado, cada punto x_0 de la condición inicial define una característica, que en el caso que hemos tratado es una recta con pendiente igual a $g(x_0)$. Por ejemplo, si tenemos una condición inicial g como la que se muestra en la figura 2a, las características correspondientes son las rectas mostradas en la figura 3a, donde puede observarse que tienen una mayor pendiente las características correspondientes a los puntos donde el valor de $g(x)$ es más alto. Estas características describen como el perfil viaja y se deforma. Por ejemplo en la figura 2b se muestra el perfil para un tiempo mayor que cero (pero cercano a cero), donde puede observarse que se ha desplazado hacia la derecha y también se ha sesgado hacia el mismo lado.

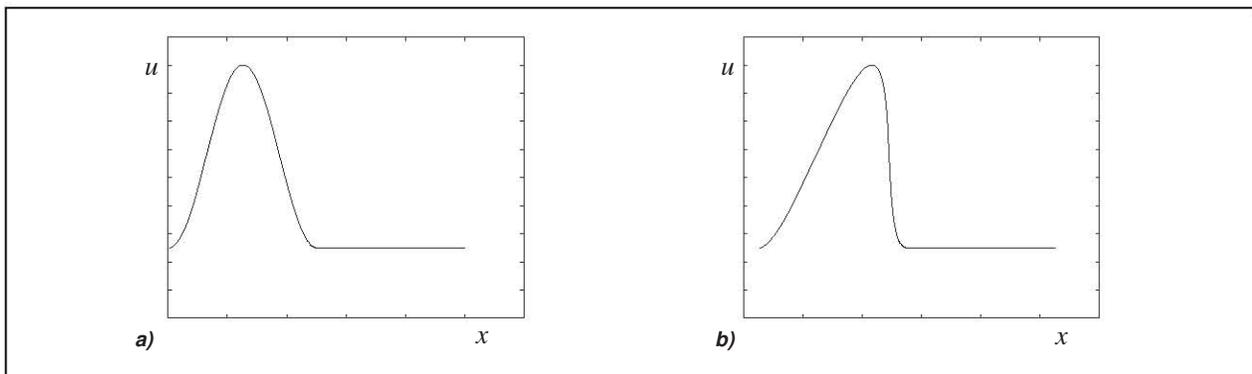


Figura 2: a) Condición inicial y b) perfil de la solución de la ecuación de Burgers no viscosa, para un tiempo mayor que, pero cercano a, cero.

En el espacio (x, t, u) cada característica define una nueva curva del tipo $(X(t), t, g(x_0))$, y el conjunto de todos los puntos sobre estas curvas forman una superficie bien definida. Sin embargo, esta superficie no es propiamente la representación de u como función de x y t porque, dado el modo como se ha construido, la superficie puede pasar por arriba de sí misma. Dicho de otro modo, como antes se ha hecho notar hay puntos en el plano (x, t) sobre los cuales pasa más de una característica. Por lo tanto, para un tiempo t suficientemente grande, la intersección de la superficie mencionada con un plano $t = \text{cte.}$ será un perfil como el que se muestra en la figura 4a. Este perfil *multivaluado* no es admisible físicamente como solución para $u(x, t)$, pero si se introduce una discontinuidad en un punto $x = s(t)$ apropiadamente escogido, puede obtenerse un perfil como el que se muestra en la figura 4b que como mostraremos a continuación, es una solución válida.

Buscaremos el punto $s(t)$ donde se encuentra la discontinuidad al tiempo t por medio de la ley de conservación

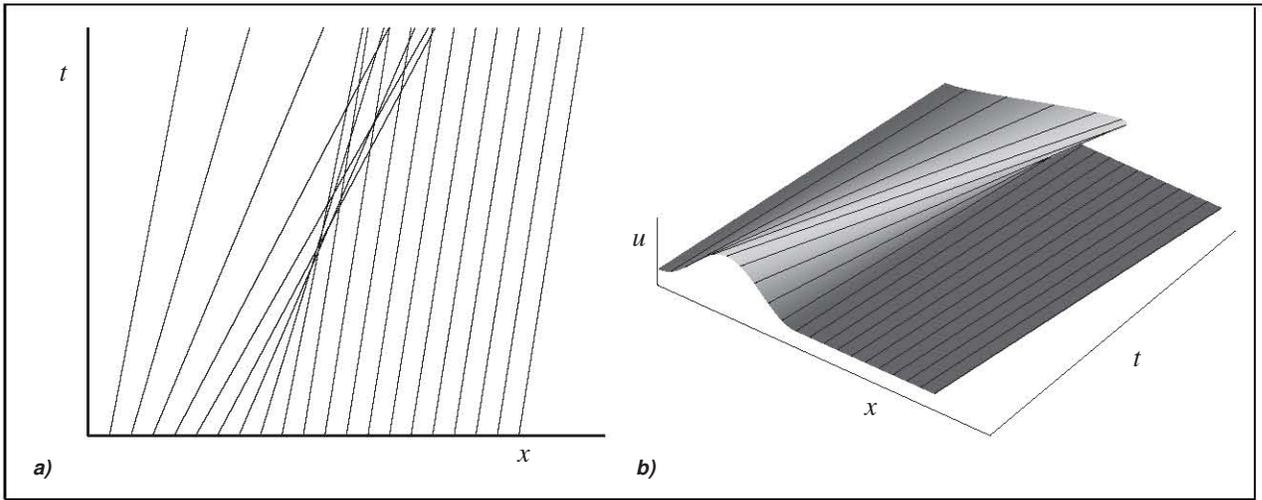


Figura 3: a) Curvas características y b) superficie formada por puntos de la forma $(X(t), t, g(x_0))$.

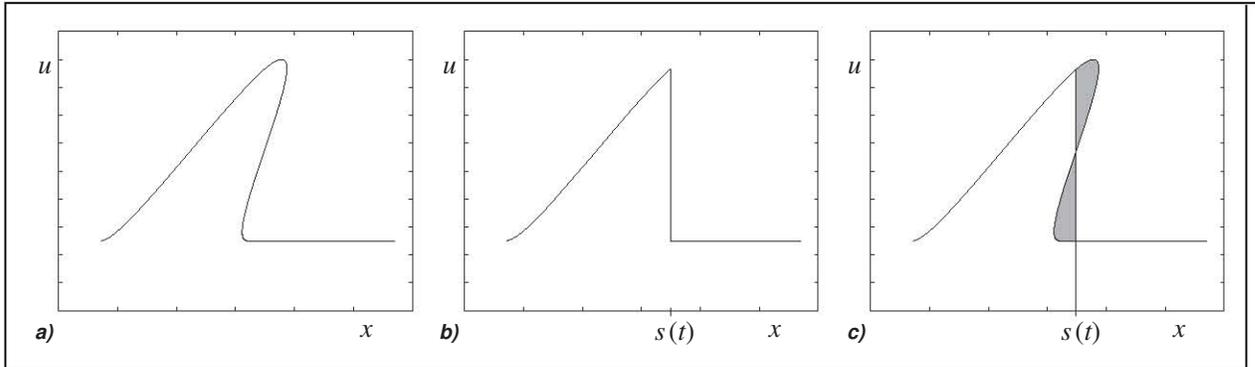


Figura 4: a) Perfil multivaluado, b) perfil con una discontinuidad en $s(t)$ y c) esquematización de la regla de las áreas iguales.

en su forma integral, evaluada en un intervalo $[x_1, x_2]$ que contenga a la discontinuidad, $x_1 < s(t) < x_2$. Entonces la ecuación (40) toma la forma

$$\frac{d}{dt} \int_{x_1}^{s(t)} \varphi \, dx + \frac{d}{dt} \int_{s(t)}^{x_2} \varphi \, dx = \tilde{q}(x_2, t) - \tilde{q}(x_1, t)$$

y por lo tanto

$$\begin{aligned} \tilde{q}(x_2, t) - \tilde{q}(x_1, t) &= \int_{x_1}^{s(t)} \varphi_t \, dx + s'(t) \varphi(s_L, t) + \int_{s(t)}^{x_2} \varphi_t \, dx - s'(t) \varphi(s_R, t) \\ &= s'(t) (\varphi(s_L, t) - \varphi(s_R, t)) + \int_{x_1}^{s(t)} \varphi_t \, dx + \int_{s(t)}^{x_2} \varphi_t \, dx, \end{aligned}$$

donde $\varphi(s_L, t)$ y $\varphi(s_R, t)$ son los límites de $\varphi(x, t)$ cuando x tiende a $s(t)$ por la izquierda y por la derecha respectivamente. Ahora, si hacemos tender $x_1, x_2 \rightarrow s(t)$ entonces las integrales del lado derecho tienden a cero y se tiene la igualdad

$$\tilde{q}(x_L, t) - \tilde{q}(x_R, t) = s'(t) (\varphi(s_L, t) - \varphi(s_R, t)),$$

donde nuevamente $\tilde{q}(x_L, t)$ y $\tilde{q}(x_R, t)$ son los límites de $\tilde{q}(x, t)$ cuando x tiende a $s(t)$ por la izquierda y por la derecha

respectivamente. A esta última ecuación se le conoce como *condición de Rankine-Hugoniot* o *condición de choque* [36], comúnmente expresada mediante

$$\Delta \tilde{q} = s'(t) \Delta \varphi . \quad (46)$$

Debido al modo como se ha construido, esta condición expresa físicamente la ley de conservación a través de la discontinuidad y es importante porque, como se ha dicho, permite encontrar la trayectoria $s(t)$ sobre la cual se propaga la discontinuidad, o bien, la onda de choque en el plano (x, t) . Normalmente se habla en plural de *las condiciones de Rankine-Hugoniot* porque en su conjunto expresan varias leyes de conservación en un medio, a través de una discontinuidad. Por ejemplo, en el caso de mecánica de fluidos, cuando expresa la conservación de masa, momento y energía a través de una superficie de discontinuidad se obtiene [2]

$$\begin{aligned} \text{masa:} & \quad \rho_L(s'(t) - u_L) = \rho_R(s'(t) - u_R) = m , \\ \text{momento:} & \quad m(u_L - u_R) = p_L - p_R , \\ \text{energía:} & \quad m(e_L + \frac{1}{2}u_L^2 - e_R - \frac{1}{2}u_R^2) = p_L u_L - p_R u_R . \end{aligned}$$

De este modo, las condiciones de Rankine-Hugoniot son expresiones discretas de las leyes de conservación y permiten proponer soluciones discontinuas, que deben satisfacer la expresión diferencial de la ley de conservación en las regiones continuas y las condiciones de Rankine-Hugoniot en las regiones discontinuas. En el caso de la ecuación de Burgers no viscosa, como $\tilde{q} = \frac{1}{2}\varphi^2$ tenemos

$$\begin{aligned} s'(t) &= \frac{\frac{1}{2}(\varphi(x_L, t))^2 - \frac{1}{2}(\varphi(x_R, t))^2}{\varphi(x_L, t) - \varphi(x_R, t)} \\ &= \frac{1}{2}(\varphi(x_L, t) + \varphi(x_R, t)) \end{aligned}$$

de modo que

$$s(t) = \frac{1}{2}(\varphi(x_L, t) + \varphi(x_R, t))t + x_d , \quad (47)$$

donde x_d es una constante de integración. Entonces la solución de la ecuación de Burgers no viscosa consta de dos partes: por un lado, para las regiones continuas, la solución por características expresada en las ecuaciones (43) y (44), y por otro lado la condición de Rankine-Hugoniot, ecuación (47), sobre el choque.

Una manera alternativa de encontrar la posición del choque es la *regla de las áreas iguales de Whitham* [36], que surge de la observación de que tanto el perfil multivaluado como el discontinuo, de los cuales pueden observarse ejemplos en las figuras 4a y 4b respectivamente, satisfacen la ley de conservación¹ [36] y por lo tanto una línea vertical que represente la discontinuidad debe cortar el perfil multivaluado de modo que los lóbulos resultantes tengan áreas iguales, como lo muestra la figura 4c.

3.3. Problema de Riemann.

Finalmente, aplicaremos la solución obtenida a una condición inicial sencilla que nos permita observar algunos aspectos típicos de una propagación no lineal. El problema de Riemann para la ecuación de Burgers no viscosa, tal

¹Abusando ligeramente de la notación (en el caso multivaluado) nos referimos a la ecuación (32).

como lo señala Toro [33], es:

$$\varphi_t + \left(\frac{\varphi^2}{2}\right)_x = 0,$$

$$\varphi(x, 0) = g(x) = \begin{cases} \varphi_L & \text{si } x < 0, \\ \varphi_R & \text{si } x > 0, \end{cases}$$

donde φ_L y φ_R son constantes, $\varphi_L \neq \varphi_R$. La condición inicial en este caso tiene una discontinuidad en $x = 0$, y es constante en las regiones que le quedan a izquierda y derecha. En las regiones donde la condición inicial es constante las características son rectas paralelas, pero hay que tomar en cuenta que la pendiente de las características a la izquierda de la discontinuidad no es igual a la pendiente de las que quedan a la derecha. Como se muestra en la figura 5, hay dos casos que serán tratados por separado y nos llevarán a resultados cualitativamente distintos.

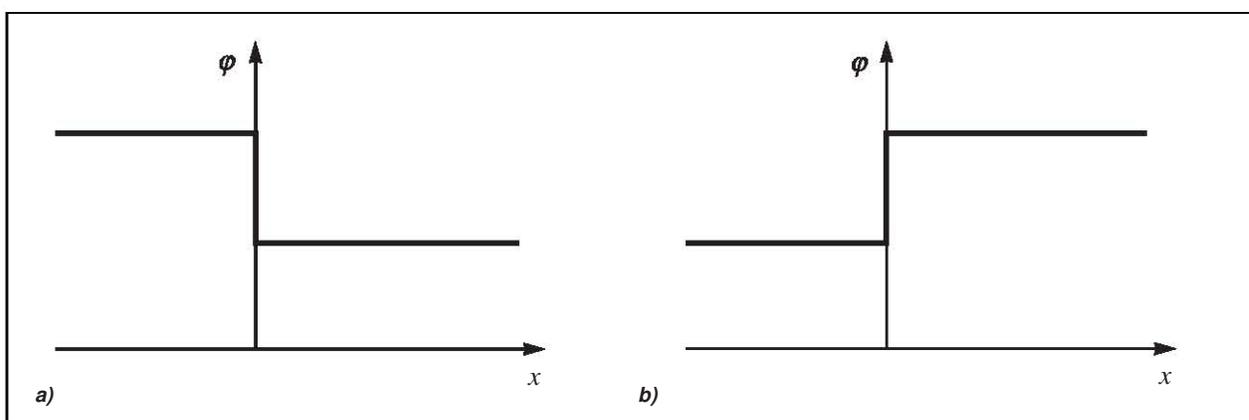


Figura 5: Condiciones iniciales correspondientes al problema de Riemann para la ecuación de Burgers no viscosa cuando **a)** $\varphi_L > \varphi_R$ y **b)** $\varphi_L < \varphi_R$.

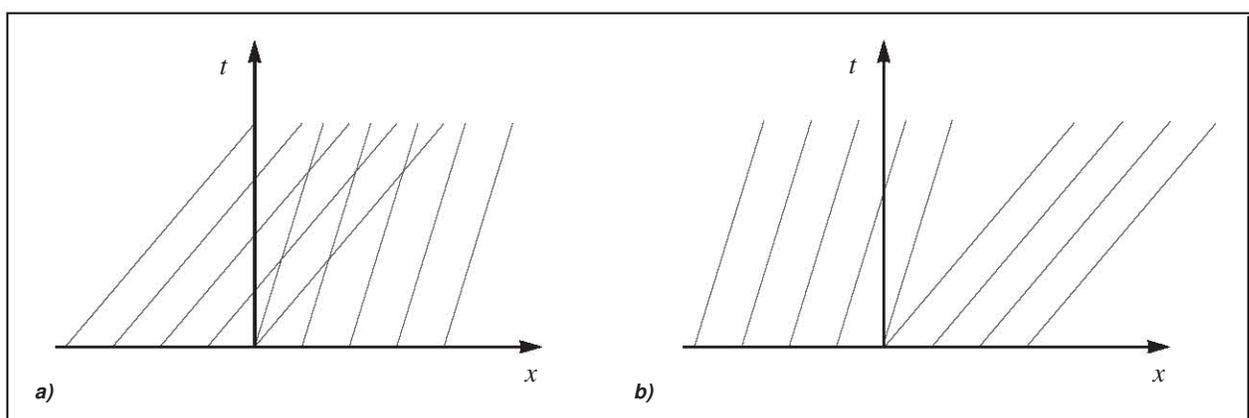


Figura 6: Características correspondientes al problema de Riemann para la ecuación de Burgers no viscosa cuando **a)** $\varphi_L > \varphi_R$ y **b)** $\varphi_L < \varphi_R$.

En el caso de que $\varphi_L > \varphi_R$ la pendiente de las características a la izquierda de la discontinuidad es la mayor y por lo tanto, como lo muestra la figura 6a, hay toda una región donde las características intersecan. Como se ha visto en la sección anterior, esto significa que existe una onda de choque y para encontrar su trayectoria $s(t)$ usaremos la

condición de Rankine-Hugoniot, que en este caso es especialmente sencilla porque los límites se calculan fácilmente:

$$\varphi(x_L, t) = \varphi_L,$$

$$\varphi(x_R, t) = \varphi_R.$$

Al sustituir estos valores en la ecuación (47) se tiene

$$s(t) = \frac{1}{2}(\varphi_L + \varphi_R)t + x_d.$$

Además, sabemos que al tiempo $t = 0$ la discontinuidad se encuentra en $s = 0$, de modo que $x_d = 0$ y

$$s(t) = \frac{1}{2}(\varphi_L + \varphi_R)t.$$

Entonces la onda de choque viaja sobre una trayectoria que es recta, parte del origen y tiene una pendiente igual al promedio de las pendientes de las características que están a sus lados izquierdo y derecho. Esta solución se muestra en la figura 7.

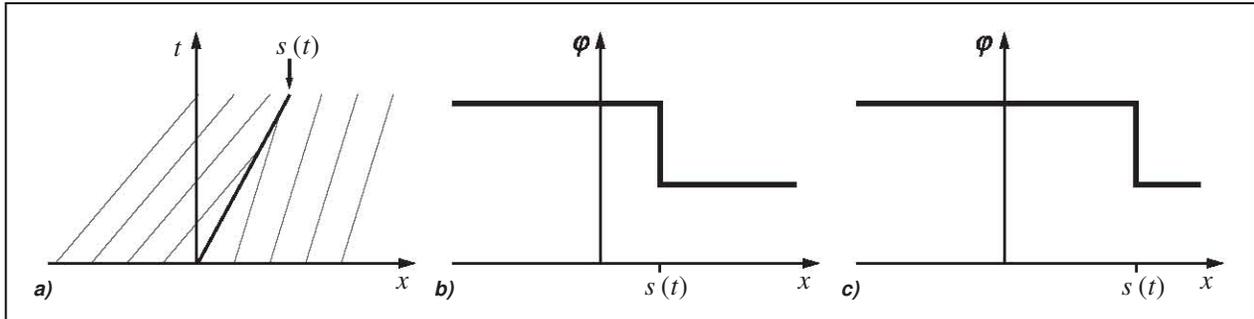


Figura 7: Solución del problema de Riemann para la ecuación de Burgers no viscosa, del caso $\varphi_L > \varphi_R$: **a)** características en el plano (x, t) , **b)** perfil de la onda de choque en el plano (x, φ) para un tiempo $t_1 > 0$, y **c)** perfil de la onda de choque en el plano (x, φ) para un tiempo $t_2 > t_1$. En todos los casos $s(t)$ denota la onda de choque.

Por otro lado, si $\varphi_L < \varphi_R$, como se muestra en la figura 6b, hay una región donde aparentemente no pasa ninguna característica y no tenemos información que nos indique el valor de φ . Sin embargo, el hecho de que las características sean líneas rectas sobre las cuales φ es constante y es igual a la pendiente de dicha recta, deriva directamente de la forma de ecuación de Burgers no viscosa, por lo que las características en esta zona también deben cumplir estas propiedades. Un juego de rectas que cumple todas estas propiedades, y la solución en este caso [33], es

$$X(t) = \varphi t,$$

donde φ varía desde φ_L hasta φ_R , este tipo de solución se llama *onda de rarefacción* y se encuentra esquematizada en la figura 8.

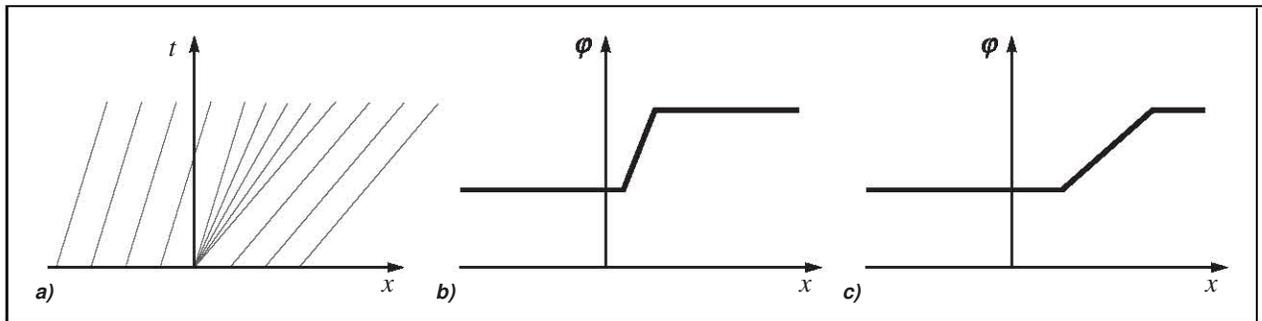


Figura 8: Solución del problema de Riemann para la ecuación de Burgers no viscosa, del caso $\varphi_L < \varphi_R$: **a)** características en el plano (x, t) , la onda de rarefacción es el abanico que surge del origen; **b)** perfil de la onda de rarefacción en el plano (x, φ) para un tiempo $t_1 > 0$; y **c)** perfil de la onda de rarefacción en el plano (x, φ) para un tiempo $t_2 > t_1$.

4. Conceptos básicos de los esquemas numéricos de diferencias finitas.

Si tenemos el problema de una ecuación diferencial, para una dimensión espacial, con valores iniciales sobre un dominio no acotado,

$$F[u] = 0 \quad (48)$$

$$u(x, t = 0) = g(x), \quad x \in \mathbb{R} \quad (49)$$

una manera de encontrar una solución aproximada es mediante la discretización del dominio, efectuando un cálculo numérico.

Aunque de manera más general las ondas se modelan usando cuatro dimensiones (tres espaciales y una temporal), en este texto consideramos únicamente propagación a lo largo de una dimensión espacial, y usaremos como base dos ecuaciones diferenciales: la primera es la ecuación de advección (17)

$$u_t + cu_x = 0,$$

donde

$$F[u] = u_t + cu_x, \quad (50)$$

como se ha descrito en la sección 2.2.1, esta ecuación es equivalente a la ecuación de onda lineal unidimensional pero con propagación en una sola dirección; y la segunda es la ecuación de Burgers no viscosa (39)

$$u_t + uu_x = 0,$$

para la cual

$$F[u] = u_t + uu_x. \quad (51)$$

Ambas ecuaciones describen perturbaciones que se mueven de izquierda a derecha.

Además, estas ecuaciones describen los casos generales siempre que el espacio sea isotropo y cada una corresponde a ciertas aproximaciones al respecto de la naturaleza de la propagación, la primera se refiere a un caso lineal y la segunda a un caso no lineal muy sencillo. Se pueden encontrar los detalles de las hipótesis que se asumen en cada caso, así como una descripción de sus soluciones en las secciones 2.2 y 3.1 respectivamente.

Un *esquema numérico de diferencias finitas* de cualquiera de estas ecuaciones diferenciales es una expresión algebraica discreta que representa a la ecuación diferencial, y donde se busca que satisfaga ciertos conceptos de convergencia, consistencia y estabilidad que serán expuestos más adelante. La ventaja que ofrece el esquema numérico es que por su carácter algebraico en la mayor parte de los casos podremos obtener una solución explícita, lo cual no necesariamente sucede con las ecuaciones diferenciales. La idea básica para obtener una *solución numérica* a partir de un esquema de este tipo es utilizar la expresión algebraica de modo que se conozcan todos los parámetros excepto uno (el valor de la solución numérica en un determinado punto), luego despejar dicho parámetro y obtener su valor; este procedimiento se puede realizar de entrada en los primeros vecinos de la condición inicial, y luego (usando los valores calculados) se puede realizar sobre el siguiente grupo de vecinos y así sucesivamente.

En las secciones siguientes se expondrán los conceptos que nos servirán para evaluar las cualidades y deficiencias

de los diferentes esquemas numéricos y la ilustración de dichos conceptos se da en la realización de los distintos esquemas, presentada en los capítulos 5 y 7.

4.1. Aproximaciones a las derivadas.

Partimos de la serie de Taylor de una función cualquiera alrededor de un punto x ; sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función con derivadas continuas hasta orden 4, entonces:

$$f(x + \Delta x) = f(x) + (\Delta x) \frac{df(x)}{dx} + (\Delta x)^2 \frac{1}{2!} \frac{d^2 f(x)}{dx^2} + (\Delta x)^3 \frac{1}{3!} \frac{d^3 f(x)}{dx^3} + O((\Delta x)^4) \quad (52)$$

donde todas las derivadas están evaluadas en el punto x , y $O((\Delta x)^4)$ indica que el residuo es de orden $(\Delta x)^4$, con Δx pequeño.

La expresión (52) es válida de igual manera para Δx tanto positivas como negativas pero aquí consideraremos Δx siempre positiva y escribiremos entonces, para un decremento,

$$f(x - \Delta x) = f(x) - (\Delta x) \frac{df(x)}{dx} + (\Delta x)^2 \frac{1}{2!} \frac{d^2 f(x)}{dx^2} - (\Delta x)^3 \frac{1}{3!} \frac{d^3 f(x)}{dx^3} + O((\Delta x)^4), \quad (53)$$

donde nuevamente las derivadas se evalúan en el punto x .

Ahora despejamos la derivada de f con respecto a x de la ecuación (52), truncando el desarrollo después del segundo término

$$\begin{aligned} \frac{df(x)}{dx} &= \frac{f(x + \Delta x) - f(x) + O((\Delta x)^2)}{\Delta x} \\ &= \frac{f(x + \Delta x) - f(x)}{\Delta x} + O(\Delta x), \end{aligned} \quad (54)$$

y del mismo modo despejamos en la ecuación (53)

$$\begin{aligned} \frac{df(x)}{dx} &= \frac{f(x - \Delta x) - f(x) + O((\Delta x)^2)}{-\Delta x} \\ &= \frac{f(x) - f(x - \Delta x)}{\Delta x} + O(\Delta x). \end{aligned} \quad (55)$$

De este modo se obtienen dos aproximaciones distintas, ambas de primer orden, para la primera derivada de una función f .

Por otro lado, si restamos la ecuación (53) de la (52) y volvemos a despejar la derivada de f con respecto a x se obtiene:

$$f(x + \Delta x) - f(x - \Delta x) = 2\Delta x \frac{df(x)}{dx} + O((\Delta x)^3),$$

y por lo tanto

$$\frac{df(x)}{dx} = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} + O((\Delta x)^2),$$

lo cual da lugar a una tercera aproximación de la primera derivada de f , pero ésta de segundo orden.

Entonces usando una serie de Taylor, valiéndonos de la función evaluada alrededor de un punto x a una distancia

Δx , ya sea antes o después de x , tenemos tres maneras de aproximar la primera derivada de f , que se muestran en el cuadro 1. Más adelante se mostrará que podemos usar otras herramientas para ampliar el rango de alternativas. Por ejemplo, en los esquemas numéricos de Lax-Friedrichs y Wendroff se utiliza el promedio entre valores vecinos.

$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x)-f(x)}{\Delta x}$	Hacia el frente
$\frac{df(x)}{dx} \approx \frac{f(x)-f(x-\Delta x)}{\Delta x}$	Hacia atrás
$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x)-f(x-\Delta x)}{2\Delta x}$	Central

Cuadro 1: Aproximaciones a la primera derivada de una función f evaluada en el punto x , obtenidas directamente de su serie de Taylor.

La aproximación central es diferente debido a que en su caso se han despreciado términos de segundo orden y superiores, mientras que en los dos primeros casos se despreció a partir de los términos de primer orden. Decimos que las aproximaciones hacia el frente y hacia atrás son de primer orden y la aproximación central es de segundo orden, y por lo tanto da lugar a una mejor aproximación.

De la misma manera es posible hacer una aproximación a la segunda derivada parcial de f , para ello sumamos la ecuación (52) a la (53), y se obtiene

$$f(x + \Delta x) + f(x - \Delta x) = 2f(x) + (\Delta x)^2 \cdot \frac{d^2 f(x)}{dx^2} + O((\Delta x)^4),$$

de donde

$$\frac{d^2 f(x)}{dx^2} = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2} + O((\Delta x)^2),$$

y entonces

$$\frac{d^2 f(x)}{dx^2} \approx \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2}.$$

Para los casos que serán expuestos más adelante, se hará referencia a una función $u = u(x, t)$, donde las aproximaciones a las derivadas parciales con respecto a x y con respecto a t siguen directamente de las expresiones del cuadro 1.

4.2. Discretización.

Discretizar un dominio significa definir un conjunto discreto de puntos aislados dentro de un dominio. El dominio que en este caso se requiere discretizar es un semiplano, $x \in \mathbb{R}$ y $t > 0$. Dicho dominio, en principio, es no acotado tanto espacial como temporalmente. Por lo tanto la discretización fácilmente podría contar con un número infinito de puntos² lo cual en términos teóricos no presenta ningún problema, pero en términos de implementación es importante recordar que son finitas tanto la cantidad de memoria como la cantidad de operaciones (sumas o multiplicaciones) que realizará efectivamente la computadora. Por lo tanto, debemos de alguna manera restringirnos a un dominio

²Incluso para dominios acotados, siempre y cuando existan puntos de acumulación, podría tenerse una discretización que conste de infinitos puntos.

acotado con una discretización finita, o bien, *ahorrar* con argumentación teórica, suficientes cálculos como para que la memoria y las operaciones necesarias sean finitas.

Una primera restricción en este sentido se impone sobre el valor de la variable temporal t , ya que estudiaremos solamente los casos donde se encuentra dentro de un intervalo $[0, T]$ con $T \in \mathbb{R}$ cuyo valor en cada caso será especificado. Esto no es suficiente para decir que el dominio es acotado porque podría no serlo respecto a la variable espacial x . Trataremos este problema de dos maneras diferentes, que se refieren a la condición inicial $g(x)$ de la ecuación diferencial. La primera de ellas consiste en tomar una condición inicial periódica, y la segunda en tomar una condición inicial con soporte compacto, es decir, que fuera de un cierto intervalo compacto (cerrado y acotado de longitud λ_s) sea idénticamente cero.

Para los casos correspondientes a las ecuaciones que usaremos (ecuación de advección y ecuación de Burgers no viscosa), una condición inicial periódica resuelve el problema porque la simetría de estas ecuaciones nos permite afirmar que el comportamiento de su solución debe ser el mismo en intervalos de longitud igual al periodo λ_p de la condición inicial $g(x)$, es decir, podemos restringir el dominio³ a un intervalo de longitud $\lambda = \lambda_p$. Y una condición inicial de soporte compacto nos permite afirmar que, suficientemente lejos del soporte de $g(x)$, la solución debe ser cero para todo tiempo en $[0, T]$, por lo tanto, nuevamente podemos restringir el dominio a un intervalo de una cierta longitud λ (en general mayor que λ_s). En la sección 4.12 se aborda este tema con mayor detalle.

Entonces el conjunto M que usaremos como discretización del dominio es un conjunto finito de puntos (x, t) que estarán dentro de un rectángulo $[x_L, x_R) \times [0, T]$, donde $\lambda = x_R - x_L$, los valores específicos de x_L y x_R (izquierda y derecha) por el momento no son importantes y será necesario determinarlos solamente al momento de hacer la implementación. Lo único que hace falta especificar es que escogeremos los puntos a intervalos regulares⁴, separados uno de otro por distancias Δx en el eje espacial y Δt en el eje temporal. Entonces

$$M = \left\{ (x, t) \in [x_L, x_R) \times [0, T] \mid x = x_L + (i - 1) \Delta x, \quad t = (j - 1) \Delta t, \quad i, j \in \mathbb{Z}^+ \right\}, \quad (56)$$

donde se ha restado 1 a los índices para que comiencen a contar desde 1 y no desde 0 (esto resulta muy conveniente porque la arquitectura del software usado para la implementación, Matlab 7, no acepta arreglos cuyo índice incluya el cero). Este conjunto se muestra esquemáticamente en la figura 9 y a menudo nos referiremos a él como *malla*.

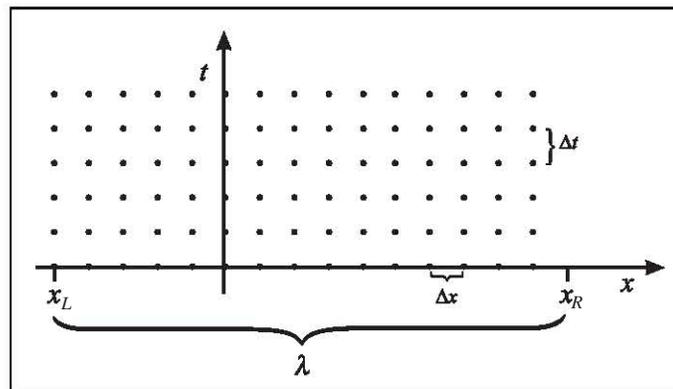


Figura 9: Discretización de un dominio rectangular a intervalos regulares.

³Tomando algunas precauciones que serán explicadas cuando sea oportuno.

⁴Se pueden desarrollar esquemas numéricos sobre discretizaciones a intervalos no regulares, lo cual resulta conveniente por ejemplo en los casos de dominios no rectangulares; sin embargo en nuestro caso no será necesario.

Para diferenciar entre $u(x, t)$ la solución exacta de la ecuación diferencial y la aproximación calculada numéricamente (la cual solamente está definida sobre los puntos de la malla M) usaremos una mayúscula para esta última, es decir, la solución numérica será denotada $U(x, t)$, y en general será más conveniente referirse a la función U por medio de los valores i, j que correspondan al punto en cuestión, de modo que en lugar de escribir

$$U(x_L + (i - 1)\Delta x, (j - 1)\Delta t)$$

se escribirá simplemente U_i^j .

En algunos casos nos referiremos a subconjuntos específicos de M como *nivel temporal* t_j . Estos subconjuntos constan de aquellos puntos (x, t) que tienen una misma $t = (j - 1)\Delta t$, en la figura 9 pueden verse como los puntos sobre una misma horizontal. Del mismo modo, es conveniente definir el vector U^j que forman los valores de U correspondientes a los puntos del nivel temporal t_j .

4.3. Esbozo de esquemas numéricos de diferencias finitas.

Una manera de formar un esquema numérico es sustituir las derivadas parciales dentro de la ecuación diferencial por aproximaciones como las que han sido descritas en la sección 4.1. Esta sustitución tiene múltiples posibilidades, las cuales nos conducirán a diferentes esquemas numéricos con diferentes características que serán estudiadas mediante los criterios que se describirán en las siguientes secciones. En este texto hemos escogido los siete esquemas que se describen en los cuadros 2 y 3 buscando que sean representativos de los diferentes tipos de esquemas. A continuación se presenta un esbozo de estos esquemas, los detalles de cada uno de ellos pueden encontrarse en las secciones 5 y 7.

Nombre	Expresión algebraica	Diagrama
Contracorriente	$\frac{U_i^{j+1} - U_i^j}{\Delta t} + c \frac{U_i^j - U_{i-1}^j}{\Delta x} = 0$	
Lax-Friedrichs	$\frac{U_i^{j+1} + \frac{1}{2}(U_{i-1}^j + U_{i+1}^j)}{\Delta t} + c \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} = 0$	
Lax-Wendroff	$U_i^{j+1} = U_i^j - \frac{1}{2} \frac{c \Delta t}{\Delta x} (U_{i+1}^j - U_{i-1}^j) + \frac{1}{2} \left(\frac{c \Delta t}{\Delta x} \right)^2 (U_{i+1}^j - 2U_i^j + U_{i-1}^j)$	
Wendroff	$\frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_i^{j+1} - U_i^j}{\Delta t} \right) + c \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right) = 0$	

Cuadro 2: Diferentes esquemas numéricos para la ecuación de advección.

■ Esquemas relativos a la ecuación de advección

1. **Contracorriente:** frecuentemente llamado *upwind* este es un esquema sencillo, obedece a un criterio de causalidad y puede obtenerse sustituyendo todas las derivadas por aproximaciones de primer orden.
2. **Lax-Friedrichs:** en este caso se usa una aproximación a segundo orden para sustituir una de las derivadas y además se usa el recurso de promediar entre puntos vecinos para evitar el problema de la inestabilidad.

Nombre	Expresión algebraica	Diagrama
Primero	$U_i^{j+1} = U_i^j - \frac{1}{2} \frac{U_i^j \Delta t}{\Delta x} (U_{i+1}^j - U_{i-1}^j) + \frac{1}{2} \left(\frac{U_i^j \Delta t}{\Delta x} \right)^2 (U_{i+1}^j - 2U_i^j + U_{i-1}^j)$	
Segundo	$\frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_{i+1}^j - U_i^j}{\Delta t} \right) + U_i^j \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right) = 0$	
Tercero	$U_i^{j+1} = U_i^j - (\Delta t) U_i^j \frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} + (\Delta t)^2 U_i^j \left(\frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} \right)^2 + \frac{(\Delta t)^2}{2} (U_i^j)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2}$	

Cuadro 3: Diferentes esquemas numéricos para la ecuación de Burgers no viscosa.

La estabilidad de un esquema numérico se describe en la sección 4.11

3. **Lax-Wendroff:** este esquema es una *mejor aproximación* a la ecuación de advección que los dos esquemas anteriores en el sentido de que es un *esquema numérico de segundo orden* mientras que los dos anteriores son esquemas numéricos de primer orden. El orden de las aproximaciones que se usan al sustituir las derivadas y el orden de un esquema numérico están relacionados, sin embargo estos son dos conceptos diferentes que se pueden ver en las secciones 4.1 y 4.10 respectivamente.
4. **Wendroff:** este también es un esquema numérico de segundo orden y se diferencia de los tres anteriores en dos aspectos: primero porque es un esquema del tipo *implícito* mientras que los tres primeros son de tipo *explícito*, y segundo porque la estabilidad de este último no está sujeta a ninguna condición.

■ Esquemas relativos a la ecuación de Burgers no viscosa

1. **Primero:** este esquema resulta de sustituir, en el esquema de Lax-Wendroff para la ecuación de advección, el valor de la velocidad c por el valor U_i^j .
2. **Segundo:** de modo similar al caso anterior, este esquema resulta de sustituir, en el esquema de Wendroff para la ecuación de advección, el valor de la velocidad c por el valor U_i^j .
3. **Tercero:** este es un esquema numérico de segundo orden con respecto a la ecuación de Burgers no viscosa (a diferencia de los dos anteriores que son de primer orden), construido por medio de un procedimiento similar al usado para construir el esquema de Lax-Wendroff.

Como se ha mencionado, los esquemas numéricos permiten encontrar una aproximación a la solución de una ecuación diferencial resolviendo sistemas de ecuaciones algebraicas. Desde luego para poder conocer el valor de U en un cierto punto mediante cualquiera de los esquemas numéricos anteriormente descritos, es necesario conocer el valor de la misma U en los puntos contiguos. En primera instancia esto lo obtendremos de las condiciones iniciales, y posteriormente se irán usando los nuevos valores de U que se vayan calculando.

4.4. Esquemas explícitos e implícitos.

Una primera diferencia que puede encontrarse en los esquemas expuestos en los cuadros 2 y 3 es que en algunos de ellos nos es posible expresar el valor de U en algún punto del nivel temporal t_{j+1} como una función de los valores de la misma U en puntos de niveles temporales anteriores. Es decir, estos esquemas involucran solamente un punto

del nivel temporal t_{j+1} , y esto en los diagramas (que se reproducen en la figura 10) puede observarse como un solo punto en el mayor nivel. A este tipo de esquemas numéricos los llamamos *esquemas explícitos* y a los casos donde

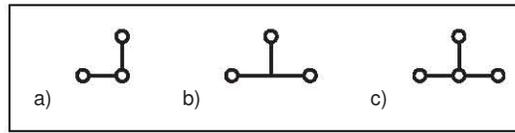


Figura 10: Diagramas para los esquemas: **a)** contracorriente, **b)** Lax-Friedrichs y **c)** Lax-Wendroff, *primero* y *tercero*.

estén involucrados dos o más puntos del nivel temporal t_{j+1} , como en el caso mostrado en la figura 11, los llamamos *esquemas implícitos*.

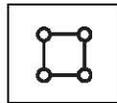


Figura 11: Diagrama para el esquema de Wendroff y esquema *segundo*.

Entre los esquemas presentados anteriormente contracorriente, Lax-Friedrichs, Lax-Wendroff, el *primero* y el *tercero* son explícitos, mientras que Wendroff y el *segundo* son implícitos.

Implementar un esquema explícito es más sencillo porque para calcular el valor de U en un determinado punto, solamente se requiere conocer el valor de U en los puntos del nivel temporal inferior, es decir, para empezar se utilizan los puntos de la condición inicial y se obtienen todos los puntos del siguiente nivel temporal, luego con los valores obtenidos se calcula otro nivel temporal más arriba y así sucesivamente. En el caso de un esquema implícito se complica el procedimiento, porque será necesario de antemano conocer el valor de U al menos en uno de los puntos contiguos al punto cuyo valor se está calculando, es decir, en el mismo nivel temporal. La solución de este problema se explica para cada uno de los esquemas implícitos en las secciones 5.5.3 y 7.4.2.

Es notable que los esquemas implícitos nos llevan, en general, a mejores resultados que los explícitos, y ello es consistente con la idea de que los esquemas explícitos utilizan solamente la información de la relación que el valor de U en un cierto punto debe guardar con el estado previo del sistema, mientras que los esquemas implícitos utilizan además la información de la relación que deben guardar los valores de U de diferentes puntos a un mismo tiempo. En la sección 4.7 se muestra como a cada esquema implícito le corresponde un esquema explícito que es equivalente.

4.5. Número CFL.

Una manera de simplificar la notación de los esquemas ya presentados en torno a la ecuación de advección es agrupar c , Δx y Δt en un solo número que resulta ser adimensional:

$$C = \frac{c \Delta t}{\Delta x}, \quad (57)$$

donde a C se le conoce como *número de Courant* o también como *número CFL* por Courant, Friedrichs y Lewy [33]. Su carácter adimensional resulta del hecho de que es un cociente de velocidades, c sobre $\frac{\Delta x}{\Delta t}$, la primera es la velocidad con la cual se propaga la solución de D'Alembert, y la segunda es una velocidad característica de la

discretización del problema, dicha es la velocidad con la cual se propaga la información en los esquemas numéricos, puede encontrarse más información sobre este punto en la sección 4.13.

Más adelante se mostrará que la exactitud con que U aproxima a u dependerá únicamente del valor de C , y no de los valores independientes de c , Δx y Δt . Muy particularmente, para los esquemas explícitos el valor de C determinará si el esquema es estable o no lo es (vea sección 4.11).

4.6. Notación polinomial.

Otra manera de escribir los diferentes esquemas numéricos que resultará útil en las secciones posteriores utiliza los operadores de traslación (shift operators) S_+ y S_- definidos por las siguientes relaciones:

$$S_+ U_i^j \equiv U_{i+1}^j, \quad (58)$$

$$S_- U_i^j \equiv U_{i-1}^j. \quad (59)$$

Es decir, cuando se aplican sobre el valor U de un determinado punto devuelven el valor U del punto espacialmente anterior o posterior, dependiendo de si se ha usado S_- o S_+ .

De este modo es posible expresar ciertos esquemas numéricos por medio de polinomios de estos mismos operadores aplicados sobre la solución numérica en diferentes niveles temporales. De manera general, para un esquema de dos niveles temporales:

$$W(S_-, S_+) U_i^{j+1} = V(S_-, S_+) U_i^j.$$

Por ejemplo, para un esquema numérico⁵ definido por

$$U_i^{j+1} = U_i^j - C(U_i^j - U_{i-1}^j),$$

donde C es el número CFL, los polinomios correspondientes V y W son:

$$W(A, B) = 1,$$

$$V(A, B) = (1 - C) + CB = 1 - C(1 - B).$$

Esto se debe a que, al aplicarlos a los operadores de traslación

$$W(S_-, S_+) = 1,$$

$$V(S_-, S_+) = 1 - C(1 - S_-),$$

y luego aplicarlos a los valores de U sobre los puntos de la malla

$$W(S_-, S_+) U_i^{j+1} = U_i^{j+1},$$

$$V(S_-, S_+) U_i^j = U_i^j - C(U_i^j - U_{i-1}^j),$$

⁵Este es el esquema numérico contracorriente, vea sección 5.1

se recupera la ecuación que define al esquema numérico:

$$W(S_-, S_+) U_i^{j+1} = V(S_-, S_+) U_i^j \quad \Longleftrightarrow \quad U_i^{j+1} = U_i^j - C(U_i^j - U_{i-1}^j) .$$

La notación polinomial ciertamente es útil para describir una amplia gama de esquema numéricos; sin embargo, para un término $(U_i^j)^2$ no existe representación alguna del tipo $V(S_-, S_+)$. De hecho no existe para ningún término que involucre productos entre factores U . Por lo tanto, esta notación no podrá aplicarse para esquemas numéricos que tengan este tipo de términos y muy particularmente para esquemas correspondientes a ecuaciones no lineales. Es posible entonces usar notación polinomial para describir esquemas que representen la ecuación de advección, pero no la ecuación de Burgers no viscosa.

Bajo esta notación la condición para que un esquema sea explícito es que cumpla con

$$W = 1 .$$

4.7. Notación matricial.

En algunos casos también podemos utilizar una notación matricial para expresar los esquemas numéricos. Primero consideraremos a los valores de U correspondientes a los puntos de un mismo nivel temporal t_j como un vector columna \mathbf{U}^j . Los esquemas relativos a la ecuación de advección se pueden siempre expresar mediante ecuaciones donde de un lado aparece una combinación lineal de los elementos de \mathbf{U}^{j+1} y del otro lado aparece una combinación lineal de los elementos de \mathbf{U}^j .

Entonces simplemente se expresan las combinaciones lineales por medio de las matrices \mathbf{A} y \mathbf{B} , que resultan ser cuadradas $i_b \times i_b$ (y prácticamente tridiagonales), es decir,

$$\mathbf{A}\mathbf{U}^{j+1} = \mathbf{B}\mathbf{U}^j .$$

La similitud de las matrices \mathbf{A} y \mathbf{B} con matrices tridiagonales deriva de que las expresiones algebraicas que definen a los esquemas numéricos que se han expuesto involucran solamente tres diferentes *niveles espaciales*, es decir, aparecen como índices inferiores de U solamente $i - 1$, i e $i + 1$. Luego, el hecho de que \mathbf{A} y \mathbf{B} no sean en todos los casos estrictamente tridiagonales por un lado se debe a que, dependiendo del esquema numérico, podrían aparecer ceros sobre alguna de las tres diagonales centrales; y por otro lado, en el caso de condiciones iniciales periódicas, como se usan los valores que corresponden a puntos de extremos opuestos de la malla, podrían aparecer valores diferentes de cero en las esquinas superior derecha o inferior izquierda.

Si el esquema es explícito debe observarse que la matriz \mathbf{A} es precisamente la matriz identidad. En el caso contrario el esquema es implícito. Sin embargo, esta notación permite además observar que debe existir un esquema numérico *explícito* que le sea equivalente. Para construirlo simplemente es necesario calcular la inversa de la matriz del lado izquierdo \mathbf{A} y multiplicarla del mismo lado izquierdo a toda la ecuación:

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{A}\mathbf{U}^{j+1} &= \mathbf{A}^{-1}\mathbf{B}\mathbf{U}^j , \\ \mathbf{U}^{j+1} &= \mathbf{A}^{-1}\mathbf{B}\mathbf{U}^j . \end{aligned}$$

Este esquema *explícito equivalente* es necesariamente diferente, porque para calcular U en cada punto a un cierto nivel temporal se requiere información al respecto de **todos** los puntos del nivel temporal anterior, mientras que los esquemas implícitos que serán expuestos y estudiados más adelante requieren solamente los valores de algunos puntos vecinos. Esto se debe a que la matriz $\mathbf{A}^{-1}\mathbf{B}$ bien puede tener, y no conocemos ninguna restricción que lo impida, todas sus entradas distintas de cero.

4.8. Monotonicidad.

Los esquemas explícitos se pueden, en general, expresar de la siguiente forma

$$U_i^{j+1} = H(U_{i-i_L}^j, \dots, U_{i+i_R}^j)$$

donde i_L e i_R son enteros no negativos (las letras L y R se refieren a izquierda y derecha). Para ecuaciones lineales la función H es de la forma

$$U_i^{j+1} = \sum_{k=i_L}^{i_R} b_k U_{i+k}^j, \quad (60)$$

donde los coeficientes b_k son reales.

De acuerdo a lo anterior podemos identificar como **monótonos** a los esquemas cuyos coeficientes b_k son no negativos, es decir,

$$b_k \geq 0, \quad \forall k. \quad (61)$$

De manera equivalente, esta condición puede también ser expresada mediante la función H , de la siguiente manera:

$$\frac{\partial H}{\partial U_k^j} \geq 0, \quad \forall k.$$

4.9. Expresiones derivadas de las ecuaciones diferenciales.

En las secciones posteriores, además de calcular series de Taylor en diferentes ocasiones, será necesario recurrir a la forma de las ecuaciones diferenciales mismas para poder simplificar un poco. Por ejemplo, si u es solución de la ecuación de advección podemos intercambiar las apariciones de u_t por $(-cu_x)$ y derivar la ecuación parcialmente con respecto a una y otra variable, para luego hacer sustituciones recursivamente⁶, y de este modo obtener algunas identidades. Del mismo modo, en el caso de la ecuación de Burgers no viscosa, siguiendo el mismo procedimiento encontramos que es válido otro grupo de identidades. A continuación se muestra una lista de las que se cumplen, en ambos casos:

- Si u satisface la ecuación de advección (17)

$$u_t + cu_x = 0$$

⁶Para ello se asume que la función es suficientemente suave como para que se pueda intercambiar el orden de las derivadas parciales.

entonces

$$\begin{aligned}
 u_t &= -cu_x \\
 u_{tx} &= u_{tx} = -cu_{xx} \\
 u_{tt} &= c^2 u_{xx} \\
 u_{txx} &= u_{xxt} = -cu_{xxx} \\
 u_{ttx} &= u_{xtt} = c^2 u_{xxx} \\
 u_{ttt} &= -c^3 u_{xxx}
 \end{aligned} \tag{62}$$

- Si u satisface la ecuación de Burgers no viscosa (39)

$$u_t + uu_x = 0$$

entonces

$$\begin{aligned}
 u_t &= -uu_x \\
 u_{tx} &= -u_x^2 - uu_{xx} \\
 u_{tt} &= 2uu_x^2 + u^2 u_{xx}
 \end{aligned} \tag{63}$$

4.10. Error de truncamiento local, consistencia y orden de un esquema numérico.

El error de truncamiento local y el orden son dos parámetros asociados con cada uno de los esquemas numéricos que nos permiten evaluar, en un sentido muy concreto, la calidad de la aproximación con respecto a la ecuación diferencial que les corresponde. Intuitivamente esperamos que cuando se refine la malla los esquemas numéricos aproximen mejor a la ecuación diferencial. Los parámetros arriba mencionados permiten establecer cuáles esquemas funcionan mejor en este sentido.

Para calcular el error de truncamiento partimos de la ecuación en diferencias finitas, y sustituimos las apariciones de U por una función cualquiera f evaluada en el punto que le corresponda, es decir, por ejemplo,

$$\begin{aligned}
 U_i^j &= f(x, t) \\
 U_{i+1}^j &= f(x + \Delta x, t) \\
 U_i^{j+1} &= f(x, t + \Delta t) \\
 &\vdots \\
 &\text{etc.}
 \end{aligned}$$

Luego desarrollamos la función f en su serie de Taylor alrededor de (x, t) y debemos obtener una expresión de la forma

$$F[f] + F_L[f] = 0 \tag{64}$$

donde F es un operador diferencial que corresponde a la ecuación diferencial que aproxima el esquema numérico (vea el inicio de esta sección 4) y el operador F_L es un operador diferencial que consiste de derivadas parciales de órdenes superiores y factores Δx y Δt .

En las secciones posteriores, al calcular el error de truncamiento local para cada uno de los esquemas numéricos, omitiremos algunos pasos; sin embargo, debe tenerse en mente la expresión (64) porque como el esquema numérico

es una ecuación algebraica, hay varias maneras equivalentes de expresarlo, pero no cualquiera nos llevará a una expresión de la forma (64). Por ejemplo, el esquema numérico

$$\frac{U_i^{j+1} - U_i^j}{\Delta t} + c \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} = 0$$

algebraicamente equivale a la expresión

$$U_i^{j+1} = \frac{c(\Delta t)}{2(\Delta x)} (U_{i+1}^j - U_{i-1}^j) + U_i^j,$$

pero esta última no sería útil porque al hacer los desarrollos de Taylor se obtiene algo de la forma

$$(f_t + cf_x) \Delta t + \dots = 0,$$

que no funciona porque aparece un factor Δt acompañando al operador $F = \frac{\partial}{\partial t} + c \frac{\partial}{\partial x}$.

La expresión (64) define una nueva ecuación diferencial por medio del operador $F + F_L$ cuya solución f en general será diferente de u , la solución de la ecuación diferencial original

$$F[u] = 0. \quad (65)$$

Por lo tanto, si sustituimos f por u en la ecuación (64), ésta se cumplirá sólo de modo aproximado; en otras palabras, del lado derecho no tendremos precisamente un cero sino un número pequeño, al cual denotaremos por L , y lo llamaremos *error de truncamiento local*. Tenemos pues

$$F[u] + F_L[u] = L, \quad (66)$$

y como u satisface la ecuación (65), también

$$L = F_L[u]. \quad (67)$$

El valor de L depende del punto en que estén evaluadas las derivadas parciales; es por esta razón que se le llama error de truncamiento *local*, ya que L es una función de x y de t .

Ahora, si bajo alguna condición es posible acotar el valor de $F_L[f]$ entonces dicha cota nos señalará la calidad de la aproximación del esquema numérico, porque cuando $F_L[f]$ es muy pequeño la ecuación diferencial obtenida del esquema numérico

$$F[f] + F_L[f] = 0 \quad (68)$$

se reduce prácticamente a la ecuación diferencial original

$$F[f] \cong 0. \quad (69)$$

Para poder encontrar esta cota, lo importante es observar las potencias con las cuales aparecen Δx y Δt dentro del operador F_L , en primer lugar porque eso es independiente del hecho de que F_L esté aplicado a u o a f , y en segundo lugar porque el refinamiento de la malla se refiere precisamente al cambio en los valores Δx y Δt .

Dicho refinamiento se puede hacer de diferentes maneras, correspondientes a la relación que exista entre Δx y

Δt ,

$$\Delta x = h(\Delta t) \quad h : \mathbb{R} \rightarrow \mathbb{R},$$

y esto se refleja en el valor de L . Incluso puede darse el caso de que L tienda a infinito cuando se refine la malla, al L contener términos del tipo $\frac{(\Delta x)^2}{\Delta t}$ cuando Δt tiende a cero con Δx constante. Este es el caso, por ejemplo, del esquema de Lax-Friedrichs. Hechas estas precisiones podemos enunciar el concepto de *consistencia*, cuya definición [14] simplemente pide que L tienda a cero cuando se refina la malla, y las observaciones arriba hechas señalan que algunos esquemas numéricos son *consistentes* con sus respectivas ecuaciones diferenciales bajo la condición de que Δx y Δt guarden cierto tipo de relación.

De la misma manera se define [30] el *orden de un esquema numérico*, u *orden de precisión* (en inglés accuracy order), como el valor $Q \in \mathbb{Z}$ más grande que cumple con

$$\|L\| = O((\Delta t)^Q) \quad (70)$$

donde los dobles barras verticales se refieren a la norma de Hilbert

$$\|L\| = \left(\int_{x_1}^{x_2} |L|^2 dx \right)^{1/2}.$$

Es importante no perder de vista que el valor de Q dependerá de la función h que se haya escogido. En nuestro caso, del mismo modo que se ha planteado en el texto de Le Veque [35], para estos fines supondremos que Δx guarda una relación proporcional con Δt de modo que h es simplemente una multiplicación por una constante. Por el modo como está definido C en la ecuación (57)

$$C = \frac{c \Delta t}{\Delta x},$$

esta suposición equivale a tomar C como una constante al momento de refinar la malla.

Por otro lado, en nuestro caso, no será necesario evaluar la integral correspondiente a la norma de Hilbert debido a que será posible acotar directamente $|L|$ mediante notación asintótica

$$|L| = O((\Delta t)^Q) \quad \forall x,$$

y por lo tanto, siempre que L esté acotado para algún valor de t , se cumple la relación (70). Eso significa que al refinar la malla bajo la restricción mencionada, el error de truncamiento L tiende a cero tan rápido como $(\Delta t)^Q$, y como se ha dicho, del mismo modo esto implica que $F_L[f]$ tiende a cero al mismo ritmo. Por lo tanto, Q es una medida de la calidad de la aproximación del esquema numérico con respecto a la ecuación diferencial; un esquema numérico representa una mejor aproximación cuando su orden Q es mayor.

4.11. Convergencia y estabilidad.

En la sección anterior se han establecido parámetros de calidad en torno a la manera en que un esquema numérico aproxima a una ecuación diferencial. Sin embargo, no se ha tratado la relación que guardan la soluciones de una y de otra, y esto es precisamente lo más importante. Cuando se trabaja con esquemas numéricos lo que se busca es aproximarse a la solución u de la ecuación diferencial por medio de la solución numérica U . Llamamos *convergente* a un esquema numérico si al refinar la malla la solución numérica U tiende en todos sus puntos a la solución exacta

u.

Sin embargo, no es una tarea fácil determinar a partir de su forma si un esquema es convergente o no lo es. Para poder hacerlo usaremos un teorema que relaciona la convergencia con el concepto de *estabilidad*. Un esquema numérico es *estable* si existe una cota superior para la magnitud de la solución numérica, que sea independiente del número de pasos temporales que se hayan dado [30]. En otras palabras el esquema es *estable* siempre que se cumpla

$$\exists z \in \mathbb{R} \quad \text{tal que} \quad |U_i^j| < z \quad \forall i, j.$$

El hecho de que un esquema numérico sea estable significa que los valores U_i^j no crecen indefinidamente a medida que se incrementa j , sino que están siempre dentro de algún rango. Pedir que un esquema sea estable es mucho menos que pedir que sea convergente; además, es más sencillo evaluar estabilidad que convergencia, por lo que resulta muy útil el *teorema de equivalencia de Lax* que, como se había mencionado, relaciona ambos conceptos.

TEOREMA DE EQUIVALENCIA DE LAX

Dado un problema de condiciones iniciales bien planteado y un esquema numérico de diferencias finitas que sea consistente con él, la estabilidad es condición necesaria y suficiente para la convergencia.

La definición de un problema de condiciones iniciales *bien planteado* así como la demostración de este teorema puede encontrarse en el texto de Richtmyer [30] y dada su formulación en términos de espacios de Banach (fuera de los alcances de este texto) aquí solamente se menciona que las ecuaciones de advección y de Burgers no viscosa bajo condiciones iniciales continuas que sean periódicas o de soporte compacto, cumplen con los requisitos de un problema bien planteado, es decir, el teorema de equivalencia de Lax es aplicable a los casos que en este texto son expuestos.

Para evaluar la estabilidad de un esquema numérico utilizaremos la condición de estabilidad de Von Neumann. El espíritu de esta prueba es entender al vector \mathbf{U}^j como una suma de funciones senoidales (componentes); es decir, por medio de su transformada discreta de Fourier establecer que ninguna de las amplitudes de estas componentes se haga arbitrariamente grande a medida que se van dando pasos temporales. La ventaja de tomar la transformada discreta de Fourier es que, como se mostrará a continuación, la amplificación de las componentes⁷ no depende de la amplitud que tengan originalmente dichas componentes.

Para enunciar formalmente la condición de estabilidad de Von Neumann primero escribimos la transformada discreta de Fourier ξ^j del vector \mathbf{U}^j , cuya definición puede consultarse por ejemplo en [3], y expresada bajo la notación del presente texto queda en la siguiente forma:

$$U_i^j = \sum_{m=1}^{i_b} \xi_m^j e^{-i \frac{2\pi m i}{i_b}}.$$

Se usa i para denotar la unidad imaginaria puesto que i ya se ha empleado como índice sobre el eje espacial. Ahora sustituimos en la ecuación correspondiente al esquema numérico usando la notación polinomial expuesta en la

⁷Con amplificación nos referimos al factor por el cual cambia la amplitud de las componentes, en otras palabras, dicha amplificación es numéricamente igual al cociente: amplitud de una componente en un determinado nivel temporal, sobre la amplitud de la misma componente en el nivel temporal anterior.

sección 4.6

$$W(S_-, S_+) \sum_{m=1}^{i_b} \xi_m^{j+1} e^{-i\frac{2\pi m i}{i_b}} = V(S_-, S_+) \sum_{m=1}^{i_b} \xi_m^j e^{-i\frac{2\pi m i}{i_b}} \quad \forall m ;$$

observamos que, en este caso, aplicar los operadores de traslación S_- y S_+ equivale a una hacer una multiplicación por $e^{-i\frac{2\pi m}{i_b}}$ y $e^{i\frac{2\pi m}{i_b}}$ respectivamente, por lo que podemos intercambiar estas operaciones para obtener

$$W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \sum_{m=1}^{i_b} \xi_m^{j+1} e^{-i\frac{2\pi m i}{i_b}} = V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \sum_{m=1}^{i_b} \xi_m^j e^{-i\frac{2\pi m i}{i_b}} .$$

Reordenando los términos se tiene

$$\sum_{m=1}^{i_b} \left[W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^{j+1} e^{-i\frac{2\pi m i}{i_b}} + V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^j e^{-i\frac{2\pi m i}{i_b}} \right] = 0 ,$$

y entonces

$$\sum_{m=1}^{i_b} \left(\left[W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^{j+1} + V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^j \right] e^{-i\frac{2\pi m i}{i_b}} \right) = 0 .$$

Ahora, como las funciones $e^{-i\frac{2\pi m i}{i_b}}$ son linealmente independientes entre sí y la expresión anterior debe cumplirse para todas las i , entonces cada uno de los coeficientes debe ser cero:

$$W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^{j+1} + V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right) \xi_m^j = 0 .$$

Finalmente, despejamos el cociente de las amplitudes,

$$\frac{\xi_m^{j+1}}{\xi_m^j} = -\frac{V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right)}{W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right)} ,$$

y observamos que la amplificación de los modos de Fourier se da solamente cuando el valor absoluto del cociente es mayor que 1. Por lo tanto la condición de estabilidad de Von Neumann puede expresarse formalmente del siguiente modo:

PRUEBA DE ESTABILIDAD DE VON NEUMANN

Sean $W(S_-, S_+)$ y $V(S_-, S_+)$ polinomios sobre los operadores de traslación S_- y S_+ , tales que describan un esquema numérico por medio de

$$W(S_-, S_+) U_i^{j+1} = V(S_-, S_+) U_i^j$$

donde $i = 1, 2, \dots, i_b$ es un índice correspondiente al eje espacial y j es otro entero correspondiente al nivel temporal. Entonces, para que el esquema sea estable, es necesario que para toda m entera, $1 \leq m \leq i_b$, se cumpla

$$\left| \frac{V\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right)}{W\left(e^{-i\frac{2\pi m}{i_b}}, e^{i\frac{2\pi m}{i_b}}\right)} \right| \leq 1 .$$

Aunque en esta última expresión no aparezca explícitamente C , el número CFL, la condición de Von Neumann es una condición precisamente para C , lo que se muestra cuando se escriben explícitamente los polinomios V y W que deben contener el factor C en algunos de sus términos.

4.12. Condiciones iniciales.

Para implementar los esquemas numéricos se pueden usar dos tipos de condiciones iniciales: periódicas o de soporte compacto, es decir, limitadas a un intervalo finito. Las funciones que quedan fuera de estas dos posibilidades son aquellas no periódicas y de dominios no acotados; sin embargo, no es posible tomarlas como base de una propagación numérica ya que, como implicarían un número infinito de puntos espaciales en cada nivel temporal, se requeriría un número infinito de operaciones en cada paso temporal.

4.12.1. Condiciones iniciales periódicas.

Cuando la condición inicial es una función periódica que se extiende por todo el eje x es posible hacer una simplificación muy importante al respecto de la implementación. Se puede hacer todo el análisis numérico dentro de un intervalo que tenga una longitud λ igual al periodo λ_p de la condición inicial $g(x)$. Esto se debe a que el valor que se obtenga para U en cada uno de los puntos de la malla depende únicamente de los valores de sus puntos vecinos. Entonces, si para dos puntos diferentes los valores de U que le corresponden a los dos juegos de vecinos respectivos son precisamente los mismos, el valor calculado de U debe ser el mismo en ambos puntos. Esto sucederá siempre que los dos puntos mencionados estén separados por una longitud igual a λ_p .

En otras palabras, la simetría de las ecuaciones que serán trabajadas, que no muestran ninguna dependencia al respecto de valores específicos de x , para condiciones iniciales periódicas implica que los resultados obtenidos de las operaciones algebraicas en cada uno de los periodos deben ser exactamente las mismas siempre y cuando el periodo sea un múltiplo de Δx . Por lo tanto podemos restringirnos al análisis de un solo periodo.

Esta estrategia de entrada presenta un problema en los extremos del intervalo de estudio: al descartar uno de sus puntos vecinos nos hace falta un dato para hacer el cálculo del siguiente nivel temporal. El problema se resuelve al usar el valor que corresponde al otro extremo, este procedimiento será explicado con detalle en la sección 5.5.2.

Para nuestro estudio las funciones periódicas que se han escogido son senoidales porque permiten aislar la propagación de cada uno de los términos de la serie de Fourier de una condición inicial más compleja. Además de que algunas ecuaciones de onda modificadas, como aquellas que se proponen en la sección 6, están resueltas a partir de una serie de términos senoidales, una condición inicial de este tipo nos facilitará la comparación entre la solución analítica de una ecuación modificada y los resultados numéricos. Además se ha escogido $[0, 1)$ como intervalo de estudio, de modo que $\lambda_p = 1$ y

$$\Delta x = \frac{1}{i_p}. \quad (71)$$

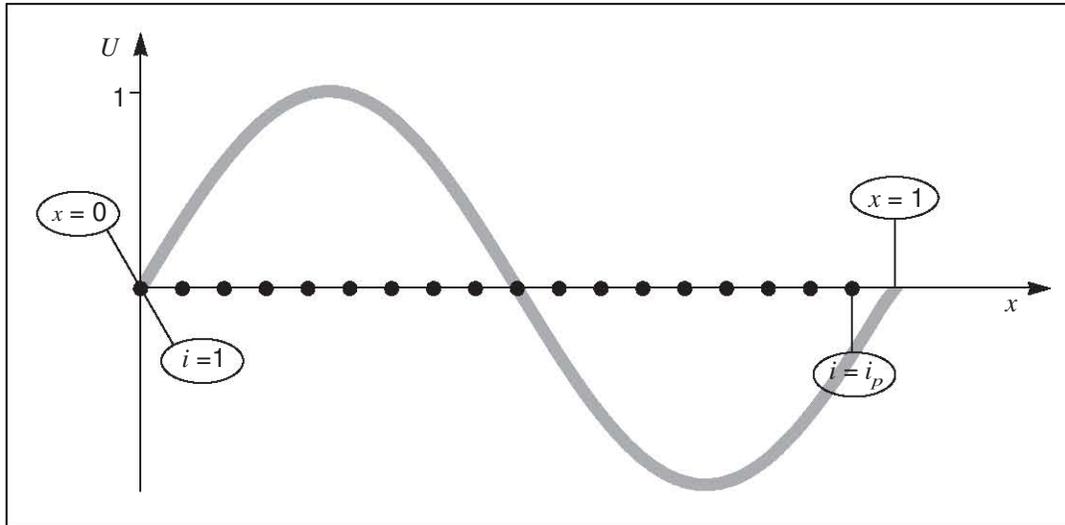


Figura 12: Condición inicial periódica.

La expresión formal de nuestra condición inicial U^1 del tipo periódico es:

$$U_i^1 = \sin\left(\frac{2\pi}{i_p}(i-1)\right) \quad \text{para } i = 1, 2, \dots, i_p. \quad (72)$$

puede observarse un diagrama en la figura 12, para un caso i_p impar.

Por consistencia con el otro tipo de condiciones iniciales es necesario especificar que en el caso periódico se cumple

$$i_b = i_p$$

donde i_b es el número de puntos espaciales involucrados en el cálculo numérico y es relevante al respecto de la estimación de la cantidad de operaciones necesarias para implementar los esquemas.

4.12.2. Condiciones iniciales de soporte compacto.

El *soporte* de una función $g(x)$ es un subconjunto de su dominio, formado por la cerradura del conjunto de puntos x tales que $g(x) \neq 0$, es decir,

$$\overline{\{x \mid g(x) \neq 0\}},$$

luego, dentro de \mathbb{R} un conjunto *compacto* es simplemente un intervalo cerrado y acotado, esta definiciones y una discusión más amplia de estos conceptos puede encontrarse por ejemplo en el texto de Lang [20]. Por lo tanto, una

función de *soporte compacto* es aquella que es idénticamente cero excepto en un determinado intervalo cerrado y acotado, dentro del cual puede anularse solamente en puntos aislados.

Este es el otro tipo de función que se usará como condición inicial $g(x)$ y adicionalmente le pediremos que sea de clase C^∞ sobre todo su dominio \mathbb{R} . La ventaja de este tipo de condición inicial es que al calcular el valor de U en un determinado punto por medio de algún esquema numérico, de los cuadros 2 o 3, solamente se hacen sumas de productos que involucran el valor de U en los puntos vecinos, por lo tanto, si U vale cero para todos los puntos vecinos involucrados el valor calculado por medio del esquema numérico debe ser precisamente cero. Entonces, el valor de U en todos los puntos que están suficientemente lejos del soporte de la condición inicial debe ser precisamente cero.

Como veremos en la sección 4.13 no podemos restringirnos a realizar el cálculo numérico y las observaciones solamente en el intervalo espacial que conforma al soporte de la condición inicial, sin embargo, es posible definir un intervalo más grande al cual sí podremos restringirnos sin ninguna pérdida de generalidad.

A este tipo de funciones definidas sobre un soporte compacto se las conoce como *pulsos*. En este texto se propone el siguiente pulso, construido a partir de un coseno, y tal que cumple con todas las condiciones antes mencionadas:

$$U_i^1 = \begin{cases} 0 & \text{si } 1 \leq i < i_a \\ \frac{1}{2} \left(1 - \cos \left(\frac{2\pi}{i_p} (i - i_a) \right) \right) & \text{si } i_a \leq i < i_p + i_a \\ 0 & \text{si } i_p + i_a \leq i \leq i_b \end{cases} .$$

Se ha elegido esta función primero porque ésta es una función suave tanto dentro del soporte como en sus extremos, es decir, ni la función ni sus derivadas presentan discontinuidades incluso en los extremos del soporte; y segundo porque el soporte resulta ser $[0, 1]$ de modo que, al igual que en el caso de condiciones iniciales periódicas

$$\Delta x = \frac{1}{i_p} . \quad (73)$$

En la figura 13 puede observarse (para un caso donde i_p es impar) la condición inicial y los puntos de la malla del primer nivel temporal, aquellos puntos que se encuentran dentro del soporte de la condición inicial se han marcado con un círculo negro ● y los que quedan fuera se han marcado con un cuadro negro ■.

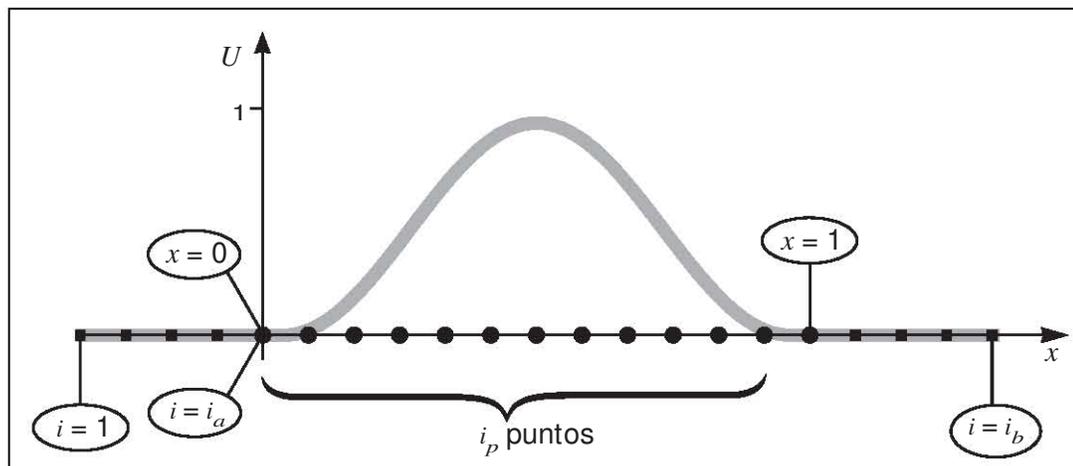


Figura 13: Condición inicial de soporte compacto.

4.12.3. Equivalencia entre ambos tipos de condiciones iniciales.

La notación matricial, descrita en la sección 4.7, nos permite observar que cada uno de los valores de \mathbf{U}^{j+1} es una combinación lineal de las componentes de \mathbf{U}^j , por lo que, si expresamos la condición inicial \mathbf{U}^1 (vector de i_b componentes) como una combinación lineal

$$\mathbf{U}^1 = a_1 \tilde{\mathbf{U}}^1 + a_2 \tilde{\tilde{\mathbf{U}}}^1,$$

donde $\tilde{\mathbf{U}}^1$ y $\tilde{\tilde{\mathbf{U}}}^1$ también son vectores de i_b componentes y los valores a_1, a_2 son escalares, entonces la solución numérica U_i^j correspondiente a \mathbf{U}^1 es la suma de las soluciones numéricas de correspondientes a $\tilde{\mathbf{U}}^1$ y $\tilde{\tilde{\mathbf{U}}}^1$, es decir:

$$U_i^j = a_1 \tilde{U}_i^j + a_2 \tilde{\tilde{U}}_i^j \quad \forall i, j. \quad (74)$$

De este modo, podemos expresar las condiciones iniciales de soporte compacto como una suma finita de funciones senoidales, es decir, condiciones iniciales periódicas. Esta descomposición se logra por medio de una transformada discreta de Fourier [3] sobre \mathbf{U}^1

$$U_i^1 = \sum_{m=1}^{i_b} \xi_m^1 e^{-i \frac{2\pi m i}{i_b}}, \quad (75)$$

donde los coeficientes ξ_m^1 son valores complejos, y el carácter finito de la suma deriva del hecho de que \mathbf{U}^1 tiene un número finito de entradas. Si se considera a \mathbf{U}^1 como un vector de i_b entradas complejas, entonces la expresión (75) equivale a un sistema lineal de $2i_b$ ecuaciones con $2i_b$ variables reales, y tendrá una solución bien definida independientemente de los valores que tenga la parte imaginaria de \mathbf{U}^1 . Es decir, podemos escoger a voluntad esta parte imaginaria y obtendremos una solución igualmente válida siempre que al interpretar el resultado sea considerada solamente la parte real, y viceversa.

En este caso es conveniente tomar la parte real desde un principio

$$U_i^1 = \sum_{m=1}^{i_b} |\xi_m^1| \sin\left(\frac{2\pi m i}{i_b} + \phi_m^1\right),$$

donde ϕ_m^1 es un parámetro real cuyo valor en este momento no será importante. La ecuación anterior se puede expresar en forma vectorial de la siguiente manera:

$$\mathbf{U}^1 = \sum_{m=1}^{i_b} |\xi_m^1| \mathbf{s}_m^1,$$

donde

$$\mathbf{s}_m^1 = \begin{pmatrix} \sin\left(\frac{2\pi m}{i_b} 1 + \phi_m^1\right) \\ \sin\left(\frac{2\pi m}{i_b} 2 + \phi_m^1\right) \\ \sin\left(\frac{2\pi m}{i_b} 3 + \phi_m^1\right) \\ \vdots \\ \sin\left(\frac{2\pi m}{i_b} i_b + \phi_m^1\right) \end{pmatrix}.$$

Los vectores $|\xi_m^1| \mathbf{s}_m^1$, que serán llamados *modos de Fourier* de \mathbf{U}^1 , son discretizaciones de funciones senoidales cuya

frecuencia depende del valor de m .

Se ha insistido en conservar el superíndice 1 en todas las variables porque la expresión (74) nos indica que si definimos el valor complejo ξ_m^j y el vector \mathbf{s}_m^j como aquellos tales que $|\xi_m^j| \mathbf{s}_m^j$ sea la solución numérica, para el nivel temporal t_j , obtenida al tomar $|\xi_m^1| \mathbf{s}_m^1$ como condición inicial, entonces debe cumplirse

$$\mathbf{U}^j = \sum_{m=1}^{i_b} |\xi_m^j| \mathbf{s}_m^j \quad \forall j.$$

Desde luego, el modo de conocer los valores específicos de $|\xi_m^j| \mathbf{s}_m^j$ es implementar el esquema numérico usando la condición inicial $|\xi_m^1| \mathbf{s}_m^1$ y esta condición inicial tiene la misma forma que las descritas en la sección 4.12.1. En este caso i_p (la cantidad de puntos del eje espacial dentro de una oscilación completa) está determinada por el valor de m y el valor i_b correspondiente a la condición inicial de soporte compacto original.

Por lo tanto, propagar una condición inicial de soporte compacto es equivalente a propagar por separado sus *modos de Fourier*, de la misma manera que se propagan las condiciones iniciales descritas en la sección 4.12.1, y luego sumar los resultados obtenidos.

4.13. Causalidad numérica.

Por el modo como se calculan los valores de U_i^n existe una relación causal *numérica* entre los puntos de la malla. A este tipo de relación causal la llamamos *numérica* para diferenciarla de aquella relacionada con la señal que describe la solución analítica de la ecuación diferencial. A excepción de los puntos correspondientes a la condición inicial, se calcula el valor de U en cada punto de la malla a partir del valor correspondiente de U en algunos de sus puntos vecinos.

Por ejemplo para un determinado esquema numérico tenemos que U_i^{n+1} guarda una relación funcional con sus vecinos del siguiente tipo:

$$U_i^{n+1} = H(U_{i-1}^n, U_i^n, U_{i+1}^n) \quad (76)$$

la cual podemos expresar gráficamente como lo muestra la figura 14.

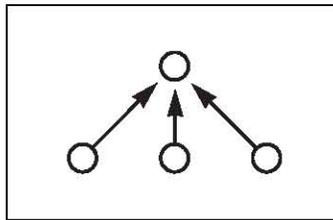


Figura 14: Relación causal numérica entre puntos vecinos determinada por un esquema numérico.

Dicha relación causal se refiere a todos los puntos de la malla, incluidos los que aparecen en la parte inferior de la figura 14. Por lo tanto los puntos que intervienen en el cálculo de un determinado U_i^n conforman una pirámide que se hace más ancha a medida que consideran niveles temporales más bajos.

Del mismo modo, habiendo escogido un determinado punto, podemos dibujar una pirámide hacia arriba con todos los puntos cuyo valor de U depende del valor del punto que hayamos escogido. Tomando como base el

ejemplo descrito en la ecuación (76), en la figura 15 se han construido estas dos pirámides usando en el primer caso una circunferencia \circ para denotar los puntos cuyo valor depende del valor de un determinado punto de la condición inicial marcado con un círculo negro \bullet ; y usando en el segundo caso una cruz delgada \times para señalar todos los puntos de cuyos valores depende el valor de un determinado punto marcado con una cruz gruesa $\mathbf{\times}$.

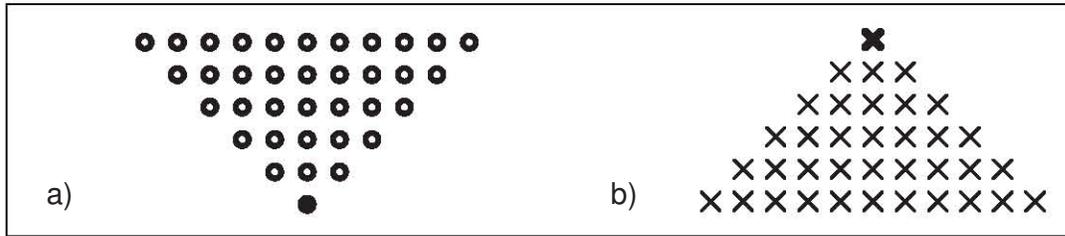


Figura 15: Diagramas de causalidad numérica: a) puntos cuyo valor de U depende valor de U en un punto en un nivel temporal anterior y b) puntos de cuyos valores de U depende el valor de U de un punto en un nivel temporal posterior.

Estas dos pirámides determinan el dominio con el cual le es posible interactuar a un determinado punto; no es posible que su valor dependa de algo que está fuera de su pirámide de cruces \times ni tampoco es posible que influya en el valor de un punto que está fuera de la pirámide de circunferencias \circ que le correspondería.

Para el caso de condiciones iniciales de soporte compacto son importantes estos diagramas porque nos permiten ahorrar una gran cantidad de cálculos.

Esto se debe por un lado al hecho de que, como se ha mencionado en la sección 4.12.2, para todos los esquemas numéricos expuestos en este texto si los valores de los cuales depende un punto son todos ceros, entonces el valor de U para dicho punto debe ser también cero. Por lo tanto, para condiciones iniciales de soporte compacto hay regiones grandes donde U vale cero en todos los puntos. Para identificar estas regiones en los diagramas se indicará con un cuadro negro \blacksquare a los puntos de la condición inicial fuera del soporte, es decir, aquellos cuyo valor de U sea necesariamente cero y se indicará con un cuadro blanco \square a los puntos cuya pirámide inferior esté formada por cuadros negros u otros cuadros blancos. Lo importante en este caso es que el valor de U en todos los puntos marcados con cuadros blancos debe ser precisamente cero y no es necesario pedir a la computadora que lo calcule. Nuevamente hemos tomado como base el ejemplo descrito en la ecuación (76) para realizar el diagrama⁸ mostrado en la figura 16.

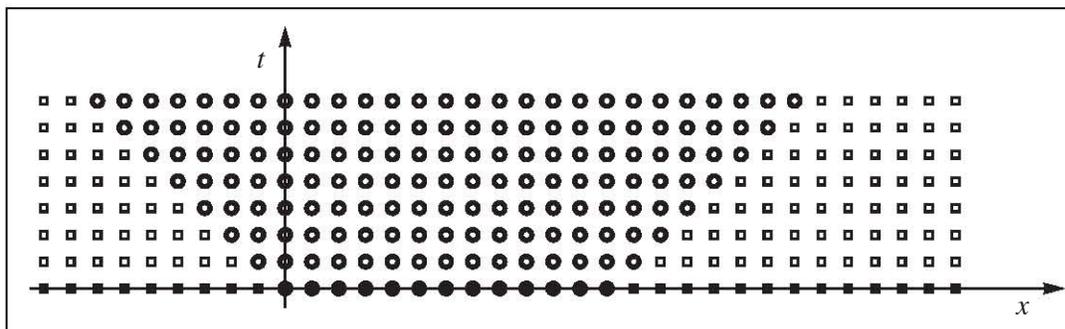


Figura 16: Primer diagrama de causalidad numérica para la relación dada por la ecuación (76).

Es posible reducir aún más la región donde han de efectuarse los cálculos. En la figura 16 se puede observar que

⁸Tanto en esta figura como en las figuras 17 y 18 se han usado valores convenientes de i_p , T y C de modo que se tengan pocos puntos y puedan realizarse fácilmente los diagramas. Como es señalado más adelante, T es el tiempo máximo que se considera en la propagación.

la región de circunferencias \bullet se ensancha a medida que avanza el tiempo. Aunque para cada esquema numérico tendremos diagramas diferentes en general sucederá que esta región resulta demasiado ancha en el sentido de que si se hace una gráfica sobre todas las circunferencias \bullet de un determinado nivel temporal observaremos que la señal ocupa solamente una pequeña parte. Por lo tanto se escoge una región sobre la cual realizar las gráficas bajo el criterio de que es en esta región donde la señal se manifiesta. En este caso el intervalo escogido está acotado en la izquierda por $x = 0$ debido a que sabemos que la señal se propaga hacia la derecha, y por el lado derecho se escoge una cota que sea un poco mayor que $x = cT$, donde T es el tiempo máximo que se considera en la propagación. En la figura 17 se ha usado el ejemplo de la ecuación (76) para construir un diagrama donde se han marcado con cruces gruesas los puntos dentro del intervalo mencionado y con cruces delgadas los puntos cuyos valores son necesarios para calcular el valor de los primeros.

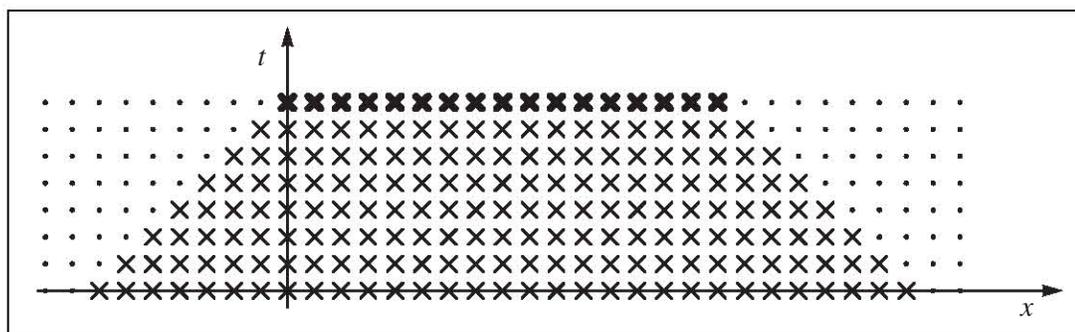


Figura 17: Segundo diagrama de causalidad numérica para el ejemplo expuesto en la ecuación (76).

El siguiente diagrama restringe la región de cálculo a los puntos donde intersectan las circunferencias \bullet con las cruces (\times o bien \times) porque los valores correspondientes al resto de las circunferencias \bullet nunca son usados ni en una gráfica, ni en el cálculo de un punto que vaya a ser graficado. Es notable que esta optimización no significa pérdida de precisión del resultado porque no se han hecho aproximaciones.

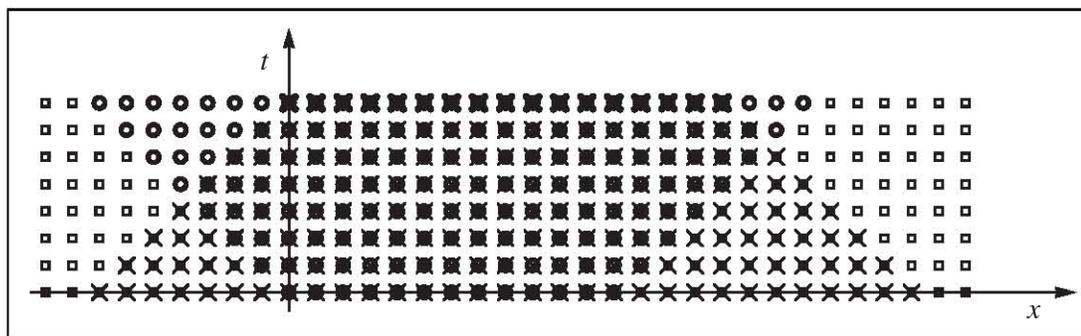


Figura 18: Superposición de primer y segundo diagrama de causalidad numérica para el ejemplo expuesto en la ecuación (76).

En la siguiente sección, en las figuras 20, 22, 24 y 26 se muestran los diagramas análogos para los esquemas numéricos particulares que hemos usado.

Para condiciones iniciales periódicas no es necesario hacer este tipo de diagramas porque no nos llevan a ninguna reducción de la región de cálculo. Como no existen puntos a los cuales corresponda un cuadro negro en la condición inicial, entonces tampoco existen cuadros blancos. Por otro lado las gráficas serán realizadas sobre el mismo periodo de la condición inicial, por lo que, la intersección entre cruces y circunferencias es toda la malla.

4.14. Dispersión numérica y difusión numérica.

En las propagaciones numéricas pueden observarse dos comportamientos que no están asociados directamente a la solución de la ecuación diferencial y por lo tanto es importante tenerlos en cuenta para procurar evitarlos. Estos son la *dispersión numérica* y la *difusión numérica*. Como su nombre lo indica estos comportamientos son análogos a la dispersión y difusión (determinadas por las propiedades físicas del medio, expuestos en la sección 2.4) pero en este caso su origen no es físico sino numérico.

Siguiendo la definición de dispersión expuesta en la sección 2.4 diremos que se observa *dispersión numérica* cuando por efecto de la aproximación numérica las soluciones numéricas del tipo senoidal se propaguen a velocidades distintas, y diremos que se observa *difusión numérica* cuando por la misma causa decrezca la amplitud de dichas perturbaciones.

En el caso de los esquemas numéricos correspondientes a la ecuación de advección es sencillo determinar el origen de dichos comportamientos observados en la solución numérica U porque su solución exacta, la solución de D'Alembert, no presenta difusión ni dispersión y es entonces posible hacer una comparación, los detalles al respecto de estas observaciones se encuentran en la sección 5.6. Por otro lado, en el caso de la ecuación de Burgers no viscosa aunque es posible obtener la solución exacta por medio de curvas características el caso requiere mayor atención. Como se ha mostrado en la sección 3, el perfil de las propagaciones (de la solución exacta) sufre una deformación que dificulta la determinación de difusión o dispersión física, y esto a su vez la determinación de difusión y dispersión numérica.

En el caso de condiciones iniciales de soporte compacto, para esquemas numéricos relativos a la ecuación de advección, será necesario aprovechar la discusión expuesta en la sección 4.12.3, es decir, pensar a los perfiles de las propagaciones como sumas de componentes senoidales y observar el cambio en sus propiedades para determinar la presencia de *difusión numérica* y/o *dispersión numérica*.

De cierto modo estos comportamientos describen el modo en que falla U al aproximar u , y de esta manera nos ayudan a saber qué tipo de errores esperar cuando se usa un esquema numérico para resolver una ecuación diferencial cuya solución exacta no se conoce. En el caso de la ecuación de advección se puede justificar teóricamente la presencia de dispersión numérica y difusión numérica mediante de las *ecuaciones modificadas* que son expuestas en la sección 6.

5. Esquemas numéricos relativos a la ecuación de advección.

5.1. Esquema contracorriente.

Un criterio para escoger el tipo de aproximaciones que usarán en el esquema numérico es la causalidad⁹. Sabemos que cualquier solución de la ecuación de advección lineal se propaga sin cambiar su perfil con velocidad constante c (solución de D'Alembert expuesta en la sección 2.2). Supongamos que c es positiva y por lo tanto la señal se propaga de izquierda a derecha. Esto implica que a un determinado punto lo que le afecta es lo que haya sucedido de su lado izquierdo en los tiempo anteriores. Por lo tanto escogemos una aproximación hacia atrás (hacia la izquierda) en el espacio y hacia el frente en el tiempo

$$\frac{U_i^{j+1} - U_i^j}{\Delta t} + c \frac{U_i^j - U_{i-1}^j}{\Delta x} = 0 \quad (77)$$

o bien, despejando U_i^{j+1} para su implementación

$$U_i^{j+1} = U_i^j - C(U_i^j - U_{i-1}^j) \quad (78)$$

donde $C = c \frac{\Delta t}{\Delta x}$ es el número CFL.

Este esquema numérico se debe a Courant, Isaacson y Rees [33], y en ocasiones puede encontrarse como esquema CIR, aunque normalmente se denomina **upwind** o **upstream** que en inglés significa *en contra del viento* o *en contra de la corriente* y en español lo podemos llamar esquema **contracorriente**. Como se ha mostrado antes, este esquema obtiene la información para calcular U en un determinado punto a partir de otros puntos que se encuentran en la dirección opuesta a la dirección en que se propaga la señal, y de ahí su nombre.

La expresión (78) determina la relación causal numérica entre puntos vecinos correspondiente a este esquema, y esta misma relación se encuentra expresada gráficamente en la figura 19. Podemos entonces construir los diagramas que nos servirán para determinar la región donde deben hacerse los cálculos en el caso de condiciones iniciales de soporte compacto; estos diagramas se encuentran en la figura 20.

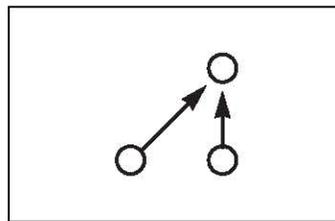


Figura 19: Relación causal numérica entre puntos vecinos para el esquema numérico contracorriente.

Por otro lado, para calcular el error de truncamiento local de este esquema numérico partimos de la expresión (77), en la cual sustituimos la solución de la ecuación de advección $u(x, t)$ y el resultado no será cero sino el error de truncamiento

$$L = \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} + c \frac{u(x, t) - u(x - \Delta x, t)}{\Delta x} . \quad (79)$$

⁹La causalidad que en este punto se describe es diferente a la causalidad numérica descrita en la sección 4.13, a la cual se hará también alusión un poco más adelante.

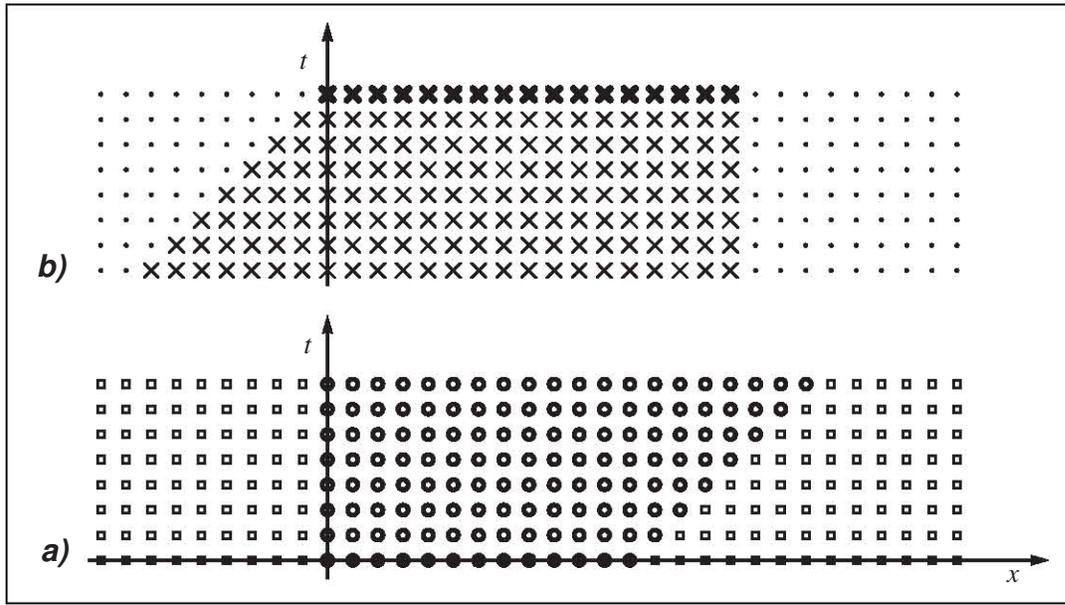


Figura 20: Diagramas de causalidad numérica para el esquema contracorriente: **a)** primero y **b)** segundo. Para la definición de esta notación vea la sección 4.13.

Ahora desarrollamos a u en serie de Taylor alrededor de (x, t) :

$$\begin{aligned}
 L &= \frac{u + (\Delta t) u_t + (\Delta t)^2 \frac{1}{2} u_{tt} + O((\Delta t)^3) - u}{\Delta t} \\
 &\quad + c \frac{u - u + (\Delta x) u_x - (\Delta x)^2 \frac{1}{2} u_{xx} + O((\Delta x)^3)}{\Delta x} \\
 &= u_t + c u_x + \Delta t \frac{1}{2} u_{tt} - \Delta x \frac{c}{2} u_{xx} + O((\Delta t)^2) + O((\Delta x)^2), \tag{80}
 \end{aligned}$$

y como u satisface

$$0 = u_t + c u_x,$$

entonces

$$L = \Delta t \frac{1}{2} u_{tt} - \Delta x \frac{c}{2} u_{xx} + O((\Delta t)^2) + O((\Delta x)^2).$$

Observamos que el esquema numérico es consistente con la ecuación de advección porque el error de truncamiento local efectivamente tiende a cero cuando tanto Δx como Δt tienden a cero. Luego, con el fin de poder aplicar la definición previamente hecha para *orden de un esquema numérico* se considera que C es una constante de modo que podemos dejar la expresión en términos de Δt substituyendo $\Delta x = \Delta t \frac{c}{C}$ y tenemos

$$\begin{aligned}
 L &= (\Delta t) \frac{1}{2} u_{tt} + O((\Delta t)^2) - \frac{c^2}{C} (\Delta t) \frac{1}{2} u_{xx} + O((\Delta t)^2) \\
 &= (\Delta t) \frac{1}{2} u_{tt} - \frac{c^2}{C} (\Delta t) \frac{1}{2} u_{xx} + O((\Delta t)^2) \\
 &= O(\Delta t),
 \end{aligned}$$

por lo que el esquema contracorriente es de primer orden.

Luego, en notación polinomial el esquema numérico contracorriente tiene la forma

$$W(S_-, S_+) U_i^{j+1} = V(S_-, S_+) U_i^j$$

donde

$$\begin{aligned} W &= 1 \\ V &= 1 - C(1 - S_-) . \end{aligned}$$

Por lo tanto la condición de estabilidad de Von Neumann toma la forma

$$\left| \frac{1 - C \left(1 - e^{-i \frac{2\pi m}{i_b}} \right)}{1} \right| \leq 1 , \quad (81)$$

y entonces

$$\begin{aligned} 1 &\geq \left| 1 - C \left(1 - e^{-i \frac{2\pi m}{i_b}} \right) \right| \\ 1 &\geq \left| 1 - C + C \cos \left(\frac{2\pi m}{i_b} \right) - iC \sin \left(\frac{2\pi m}{i_b} \right) \right| \\ 1 &\geq \sqrt{\left(1 - C + C \cos \left(\frac{2\pi m}{i_b} \right) \right)^2 + \left(C \sin \left(\frac{2\pi m}{i_b} \right) \right)^2} \\ 1 &\geq (1 - C)^2 + 2(1 - C)C \cos \left(\frac{2\pi m}{i_b} \right) + C^2 \\ 0 &\geq -2C(1 - C) + 2C(1 - C) \cos \left(\frac{2\pi m}{i_b} \right) \\ 0 &\geq C(1 - C) \left(-1 + \cos \left(\frac{2\pi m}{i_b} \right) \right) . \end{aligned}$$

Como el tercer factor es siempre negativo, entonces C y $(1 - C)$ deben tener signos iguales, por lo que la condición de Von Neumann será satisfecha solamente en el caso

$$0 \leq C \leq 1 .$$

En otras palabras el esquema contracorriente es estable, y consecuentemente convergente, cuando $0 \leq C \leq 1$. Este resultado es consistente con el hecho de que la velocidad con la cual se propaga la información en la solución numérica $\frac{\Delta x}{\Delta t}$ no puede ser mayor que la velocidad c con la cual viaja la información en la solución de D'Alembert. Es decir, la causalidad numérica no puede contradecir a la causalidad física.

Otra manera de expresar este esquema es mediante una suma del tipo expresado en la ecuación (60) donde los coeficientes resultan ser

$$b_{-1} = C, \quad b_0 = 1 - C, \quad b_1 = 0$$

y debido a que este esquema es estable solamente cuando $0 \leq C \leq 1$, se tiene que todos los coeficientes b_k son no negativos, y entonces el esquema contracorriente es monótono.

Se obtienen resultados completamente análogos cuando de antemano se asume $c < 0$ y se usa una aproximación hacia el frente (es decir, hacia la derecha) para la derivada espacial. Esto es consistente con la argumentación causal que se ha expuesto porque se cambia la dirección en que viaja la señal y se cambia del mismo modo el lado de donde se recoge la información para calcular U . Por el contrario, si se recoge información de los puntos que están al mismo lado que se desplaza la señal (ya sea que $c > 0$ o que $c < 0$) se obtiene un esquema denominado *downwind* (en español sería *a favor de la corriente*), que bajo una prueba de estabilidad de Von Neumann resulta ser incondicionalmente inestable; es decir, no representa un esquema mediante el cual se pueda resolver numéricamente la ecuación de advección.

Finalmente cabe resaltar que el caso $C = 1$ ofrece algunos resultados especiales. Primero, en la ecuación (78) nos permite observar que cuando $C = 1$ el esquema se reduce sencillamente a recorrer los valores de los puntos hacia la derecha, es decir,

$$U_i^{j+1} = U_{i-1}^j . \quad (82)$$

Para este caso, la solución de D'Alembert u debe tener el mismo valor para el punto de la malla que corresponde al punto $i - 1, j$ y para el punto que corresponde a $i, j + 1$, porque como

$$1 = C = c \frac{\Delta t}{\Delta x} ,$$

entonces

$$\frac{\Delta x}{\Delta t} = c , \quad (83)$$

y por lo tanto el punto $i - 1, j$ y el punto $i, j + 1$ se encuentran sobre una misma curva característica $X(t)$ debido a que, como se ha señalado en la sección 2.2.3, para la ecuación de advección estas curvas solamente están sujetas a la restricción

$$X'(t) = c .$$

Entonces, para el caso $C = 1$ la solución numérica es, en efecto, idéntica a la solución exacta:

$$U_i^j = u(x_L + (i - 1) \Delta x, (j - 1) \Delta t) , \quad (84)$$

donde se ha usado la expresión (56) para recuperar los valores x y t a los cuales corresponde el punto i, j .

Además, si uno sustituye la ecuación (83) en la expresión (79), correspondiente al error de truncamiento local, se tiene:

$$L = \frac{1}{\Delta t} (u(x, t + \Delta t) - u(x - c\Delta t, t)) , \quad (85)$$

y, al desarrollar en serie de Taylor

$$L = \frac{1}{\Delta t} \left(\sum_{k=0}^{\infty} \frac{1}{k!} (-c\Delta t)^k u_{x^k} - \sum_{k=0}^{\infty} \frac{1}{k!} (\Delta t)^k u_{t^k} \right) ,$$

donde se puede demostrar por inducción que $u_{t^k} = (-c)^k u_{x^k}$, y por lo tanto, como podía esperarse, cuando $C = 1$ se tiene $L = 0$.

5.2. Esquema de Lax-Friedrichs.

Al sustituir las derivadas parciales en la ecuación de advección tenemos que la aproximación central es de orden mayor que las aproximaciones hacia el frente y hacia atrás como ya se vio en el capítulo 4.1, por lo que debería brindar mejores resultados. Siguiendo este razonamiento podemos escoger aproximaciones centrales para generar un esquema numérico. Sin embargo, esta elección (llamada esquema numérico *leapfrog*¹⁰), presenta un inconveniente: se involucra información de tres distintos niveles temporales, y eso significa que como condición inicial para poder propagar la solución numérica se necesitarían conocer dos niveles temporales completos.

Para remediar este problema cambiamos la aproximación central al respecto del tiempo por una aproximación hacia el frente. Este nuevo esquema tiene también un problema: si aplicamos una prueba de Von Neumann observamos que resulta ser inestable. Sin embargo, podemos hacer otra corrección que nos permite conservar la aproximación a segundo orden de la derivada espacial y efectivamente resuelve el problema. Esta corrección consiste en cambiar U_i^j por un promedio al respecto de los puntos que están a sus lados derecho e izquierdo

$$U_i^j \approx \frac{U_{i-1}^j + U_{i+1}^j}{2}.$$

Entonces nuestro esquema numérico tiene la siguiente forma

$$\frac{U_i^{j+1} + \frac{1}{2}(U_{i-1}^j + U_{i+1}^j)}{\Delta t} + c \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} = 0, \quad (86)$$

o bien, despejando U_i^{j+1} para su implementación

$$\begin{aligned} U_i^{j+1} &= \frac{1}{2}(U_{i-1}^j + U_{i+1}^j) + C \frac{1}{2}(U_{i+1}^j - U_{i-1}^j) \\ &= U_{i-1}^j (1 + C) \frac{1}{2} + U_{i+1}^j (1 - C) \frac{1}{2}. \end{aligned} \quad (87)$$

Este esquema numérico se conoce con el nombre de **Lax-Friedrichs**. Es de señalarse que la aproximación que se usa para la derivada temporal no se deduce directamente de la serie de Taylor.

La expresión (87) determina la relación causal numérica entre puntos vecinos, misma que se ha expresado gráficamente en la figura 21. Entonces, para este esquema numérico, los diagramas que se han expuesto en la sección 4.13 resultan ser los que se muestran en la figura 22. Como se ha mencionado antes, a través de estos diagramas discriminamos los puntos de la malla donde no es necesario hacer cálculo numérico.

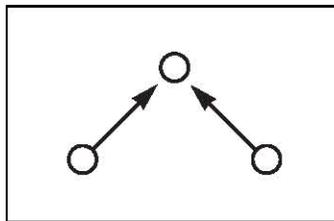


Figura 21: Relación causal numérica entre puntos vecinos para el esquema de Lax-Friedrichs.

¹⁰Cuya traducción al español es literalmente *salto de rana*, pero comúnmente esta palabra se refiere al hecho de que se ha saltado pasando sobre algo, en el caso del esquema *leapfrog* sobre un punto que es omitido.

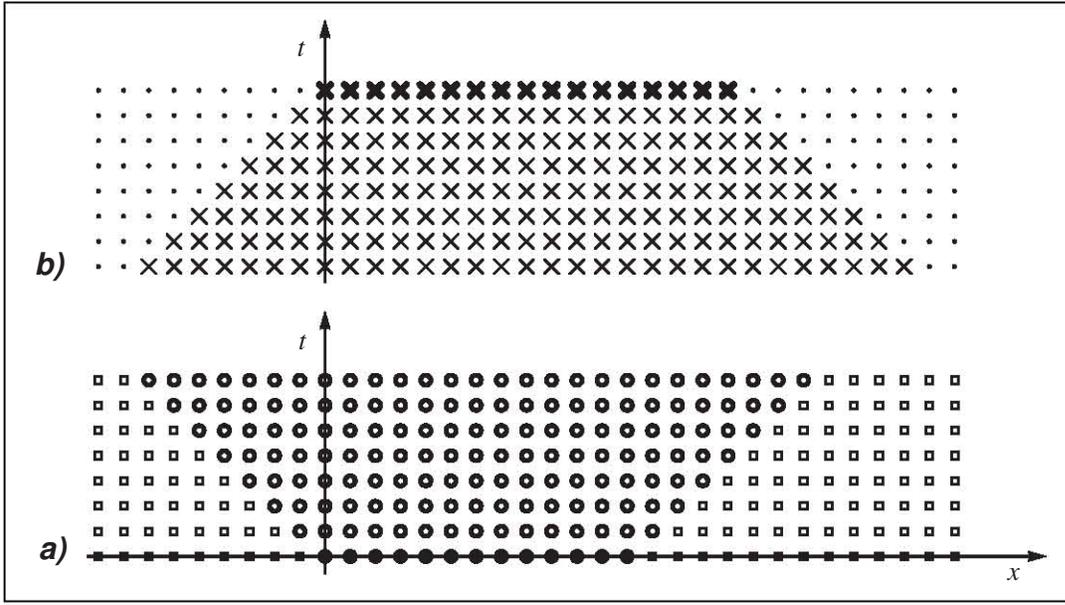


Figura 22: Diagramas de causalidad numérica para el esquema de Lax-Friedrichs: a) primero y b) segundo. Para la definición de esta notación vea la sección 4.13.

Ahora, para calcular el error de truncamiento partimos de la expresión (86)

$$0 = \frac{U_i^{j+1} - \frac{1}{2}(U_{i-1}^j + U_{i+1}^j)}{\Delta t} + c \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)}.$$

Luego, cambiamos U por la solución de la ecuación de advección y esta nueva expresión no será exactamente igual a cero, sino al error de truncamiento L :

$$L = \frac{u(x, t + \Delta t) - \frac{1}{2}(u(x - \Delta x, t) + u(x + \Delta x, t))}{\Delta t} + c \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2(\Delta x)}. \quad (88)$$

nuevamente desarrollamos por medio de su serie de Taylor los casos que no están evaluados en (x, t)

$$\begin{aligned} L &= \frac{u + (\Delta t)u_t + \frac{1}{2}(\Delta t)^2 u_{tt} + O((\Delta t)^3)}{\Delta t} \\ &\quad - \frac{\frac{1}{2}(u - (\Delta x)u_x + \frac{1}{2}(\Delta x)^2 u_{xx} + O((\Delta x)^3)) + u + (\Delta x)u_x + \frac{1}{2}(\Delta x)^2 u_{xx} + O((\Delta x)^3)}{\Delta t} \\ &\quad + c \frac{u + (\Delta x)u_x + \frac{1}{2}u_{xx} + O((\Delta x)^3) - u + (\Delta x)u_x - \frac{1}{2}u_{xx} + O((\Delta x)^3)}{2(\Delta x)} \\ &= u_t + \frac{1}{2}(\Delta t)u_{tt} + O((\Delta t)^2) - \frac{1}{2(\Delta t)}((\Delta x)^2 u_{xx} + O((\Delta x)^3)) + c(u_x + O((\Delta x)^2)) \\ &= u_t + cu_x + \frac{1}{2}(\Delta t)u_{tt} - \frac{(\Delta x)^2}{2(\Delta t)}u_{xx} + O((\Delta t)^2) + \frac{1}{\Delta t}O((\Delta x)^3) + O((\Delta x)^2), \end{aligned} \quad (89)$$

como u es solución de la ecuación de advección los dos primeros términos se cancelan y se obtiene

$$L = \frac{1}{2}(\Delta t)u_{tt} - \frac{(\Delta x)^2}{2(\Delta t)}u_{xx} + O((\Delta t)^2) + O((\Delta x)^2) + \frac{1}{\Delta t}O((\Delta x)^3).$$

Observamos que pedir que Δx y Δt tiendan a cero no es suficiente para garantizar que el error de truncamiento tienda a cero. Para que esto suceda es necesario además que el cociente $\frac{(\Delta x)^2}{\Delta t}$ tienda a cero. Por lo tanto este esquema es consistente con la ecuación de advección bajo la condición de que la función $\Delta x = h(\Delta t)$ sea tal que $\frac{(\Delta x)^2}{\Delta t}$ tienda a cero cuando Δt tiende a cero.

Luego, para conocer el orden del esquema suponemos, como antes, que C es una constante y de este modo podemos dejar la expresión en términos de Δt sustituyendo Δx por $\Delta t \frac{c}{C}$

$$\begin{aligned} L &= \frac{1}{2} (\Delta t) u_{tt} - \frac{(\Delta t) c^2}{2C^2} u_{xx} + O((\Delta t)^2) + O((\Delta t)^2) + O((\Delta t)^2) \\ &= \frac{1}{2} (\Delta t) u_{tt} - \frac{(\Delta t) c^2}{2C^2} u_{xx} + O((\Delta t)^2) \\ &= O(\Delta t) , \end{aligned}$$

y vemos entonces que el esquema de Lax-Friedrichs es de primer orden.

De la ecuación (87) observamos que los polinomios correspondientes a este esquema numérico son

$$\begin{aligned} W &= 1 \\ V &= S_- (1 + C) \frac{1}{2} + S_+ (1 - C) \frac{1}{2} \end{aligned}$$

por lo tanto la condición de estabilidad de Von Neumann toma la forma

$$\left| \frac{e^{-i\frac{2\pi m}{i_b}} (1 + C) \frac{1}{2} + e^{i\frac{2\pi m}{i_b}} (1 - C) \frac{1}{2}}{1} \right| \leq 1 , \quad (90)$$

y entonces

$$\begin{aligned} 1 &\geq \left| e^{-i\frac{2\pi m}{i_b}} (1 + C) \frac{1}{2} + e^{i\frac{2\pi m}{i_b}} (1 - C) \frac{1}{2} \right| \\ 1 &\geq \left| \frac{1}{2} (1 + C) \left(\cos\left(\frac{2\pi m}{i_b}\right) - i \sin\left(\frac{2\pi m}{i_b}\right) \right) + \frac{1}{2} (1 - C) \left(\cos\left(\frac{2\pi m}{i_b}\right) + i \sin\left(\frac{2\pi m}{i_b}\right) \right) \right| \\ 1 &\geq \left| \cos\left(\frac{2\pi m}{i_b}\right) - iC \sin\left(\frac{2\pi m}{i_b}\right) \right| \\ 1 &\geq \sqrt{\cos^2\left(\frac{2\pi m}{i_b}\right) + C^2 \sin^2\left(\frac{2\pi m}{i_b}\right)} \\ 1 &\geq \cos^2\left(\frac{2\pi m}{i_b}\right) + C^2 \left(1 - \cos^2\left(\frac{2\pi m}{i_b}\right)\right) \\ 0 &\geq (C^2 - 1) \left(1 - \cos^2\left(\frac{2\pi m}{i_b}\right)\right) . \end{aligned}$$

Como el segundo factor siempre es positivo entonces $(C^2 - 1)$ debe ser negativo, y por lo tanto el sistema será estable según la condición de Von Neumann solamente cuando

$$0 \leq |C| \leq 1 .$$

Al igual que en el esquema anterior, para el esquema de Lax-Friedrichs los casos que son señalados como inestables, son precisamente aquellos donde la velocidad con la cual se propaga la información en el esquema numérico supera a la velocidad con la cual se propaga la solución de D'Alembert, es decir, aquellos donde la causalidad numérica contradice a la causalidad física.

Por otro lado, al expresar este esquema como una suma del tipo expuesto en la ecuación (60) los coeficientes resultan ser

$$b_{-1} = \frac{1}{2}(1 + C), \quad b_0 = 0, \quad b_1 = \frac{1}{2}(1 - C).$$

Luego, como este esquema puede ser usado solamente cuando $0 \leq |C| \leq 1$, puesto que en cualquier otro caso es inestable, y dado que para dicha restricción los coeficientes b_k son no negativos, entonces el esquema es efectivamente monótono.

Finalmente, de la expresión (87) observamos que cuando $C = 1$, al igual que el esquema anterior, el esquema de Lax-Friedrichs adopta una forma particularmente sencilla

$$U_i^{j+1} = U_{i-1}^j, \quad (91)$$

es decir, al avanzar un nivel temporal los valores de los puntos simplemente se recorren un lugar hacia la derecha. Como se ha argumentado en la sección anterior, esto implica que para este caso la solución numérica U es idéntica a la solución exacta u , ecuación (84). Y del mismo modo bajo la restricción $C = 1$ la expresión (88) adquiere la forma

$$L = \frac{1}{\Delta t} (u(x, t + \Delta t) - u(x - c\Delta t, t)),$$

que implica $L = 0$, como se ha mostrado también en la sección anterior. Es decir, cuando $C = 1$, como podía esperarse, el error de truncamiento local de este esquema numérico es nulo.

5.3. Esquema de Lax-Wendroff.

El esquema de Lax-Wendroff se construye de una manera diferente a la expuesta para los dos esquemas anteriores. Se parte de dos premisas: se busca formar un esquema numérico que sea de segundo orden y se quiere que el esquema numérico sea explícito. Entonces, como queremos que nuestro esquema sea explícito, es necesario conocer el valor de $u(x, t + \Delta t)$ en términos de algo que esté evaluado en el nivel temporal anterior, t_n . Para lograrlo desarrollamos $u(x, t + \Delta t)$ en su serie de Taylor alrededor de (x, t) ,

$$u(x, t + \Delta t) = u + (\Delta t) u_t + (\Delta t)^2 \frac{1}{2} u_{tt} + O((\Delta t)^3), \quad (92)$$

luego incorporamos la información correspondiente a la ecuación que pretende aproximar nuestro esquema numérico, mediante, en este caso, la ecuación de advección,

$$u_t = -cu_x, \quad (93)$$

observando que además deben cumplirse las relaciones (62). Al sustituir (62) y (93) en la ecuación (92) se obtiene

$$u(x, t + \Delta t) = u - c(\Delta t) u_x + c^2 (\Delta t)^2 \frac{1}{2} u_{xx} + O((\Delta t)^3).$$

Ahora construimos este esquema numérico intercambiando las derivadas parciales con respecto a x por aproximaciones centrales y despreciando términos de tercer orden. Como veremos más adelante, esta elección implica efectivamente que este esquema resulta ser de segundo orden:

$$\begin{aligned}
 U_i^{j+1} &= U_i^j - c(\Delta t) \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} + c^2(\Delta t)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{2(\Delta x)^2} \\
 &= U_i^j - \frac{C}{2}(U_{i+1}^j - U_{i-1}^j) + \frac{C^2}{2}(U_{i+1}^j - 2U_i^j + U_{i-1}^j). \tag{94}
 \end{aligned}$$

Esta última expresión determina la relación causal que guardan los puntos vecinos dentro de este esquema; la figura 23 muestra gráficamente dicha relación. En la figura 24 se encuentran los diagramas, del tipo expuesto en la sección 4.13, para el esquema de Lax-Wendroff, mediante las cuales podemos discriminar los puntos de la malla donde no es necesario hacer cálculos numéricos.

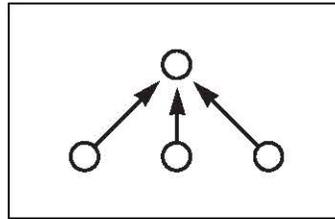


Figura 23: Relación causal numérica entre puntos vecinos para el esquema de Lax-Wendroff.

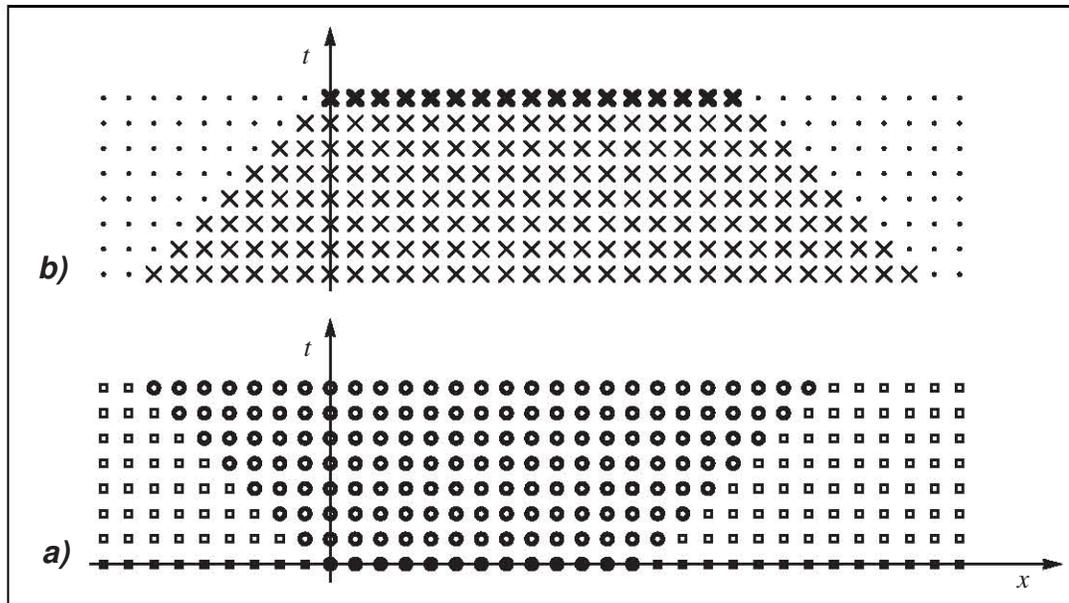


Figura 24: Diagramas de causalidad numérica para el esquema de Lax-Wendroff: a) primero y b) segundo. Para la definición de esta notación vea la sección 4.13.

Para calcular el error de truncamiento local L reordenamos la expresión (94) para tener

$$0 = \frac{U_i^{j+1} - U_i^j}{\Delta t} + c \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} - \frac{c^2(\Delta t)}{2} \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2}.$$

Sustituimos entonces U por u al modo que se ha descrito en la sección 4.10,

$$L = \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} + c \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2(\Delta x)} - \frac{c^2(\Delta t)}{2} \cdot \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2}, \quad (95)$$

y ahora hacemos el desarrollo de Taylor para cada uno de los términos:

$$\begin{aligned} \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} &= u_t + \Delta t \frac{1}{2} u_{tt} + (\Delta t)^2 \frac{1}{6} u_{ttt} + O((\Delta t)^3), \\ \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2(\Delta x)} &= u_x + (\Delta x)^2 \frac{1}{6} u_{xxx} + O((\Delta x)^3), \\ \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2} &= u_{xx} + O((\Delta x)^2). \end{aligned}$$

Por lo tanto,

$$\begin{aligned} L &= u_t + \Delta t \frac{1}{2} u_{tt} + (\Delta t)^2 \frac{1}{6} u_{ttt} + O((\Delta t)^3) \\ &\quad + c \left(u_x + (\Delta x)^2 \frac{1}{6} u_{xxx} + O((\Delta x)^3) \right) \\ &\quad - \frac{c^2(\Delta t)}{2} (u_{xx} + O((\Delta x)^2)) \\ &= u_t + cu_x + \Delta t \frac{1}{2} u_{tt} - \frac{c^2(\Delta t)}{2} u_{xx} + (\Delta t)^2 \frac{1}{6} u_{ttt} + c(\Delta x)^2 \frac{1}{6} u_{xxx} \\ &\quad + O((\Delta t)^3) + O((\Delta x)^3) + (\Delta t) O((\Delta x)^2). \end{aligned} \quad (96)$$

Debido a que u es solución de la ecuación de advección es posible usar las relaciones (62) para eliminar cuatro términos y finalmente obtener

$$\begin{aligned} L &= O((\Delta t)^2) + c O((\Delta x)^2) - \frac{c^2(\Delta t)}{2} O((\Delta x)^2) \\ &= O((\Delta t)^2) + O((\Delta x)^2) - (\Delta t) O((\Delta x)^2). \end{aligned}$$

Observamos nuevamente que cuando Δx y Δt tienden a cero L también tiende a cero, por lo que este esquema numérico es consistente con la ecuación de advección. Por otro lado también se observa que el esquema es de segundo orden debido a que al considerar $C = c \frac{\Delta t}{\Delta x}$ como una constante y sustituir Δx en términos de Δt se obtiene

$$\begin{aligned} L &= O((\Delta t)^2) + O((\Delta t)^2) - (\Delta t) O((\Delta t)^2) \\ &= O((\Delta t)^2). \end{aligned}$$

Puede observarse, a partir de la expresión (94) que la notación polinomial correspondiente a este esquema es:

$$\begin{aligned} W &= 1, \\ V &= 1 - \frac{C}{2} (S_+ - S_-) + \frac{C^2}{2} (S_+ - 2 + S_-), \end{aligned}$$

de modo que la condición de estabilidad de Von Neumann se puede expresar como

$$\left| \frac{1 - \frac{C}{2} \left(e^{i\frac{2\pi m}{i_b}} - e^{-i\frac{2\pi m}{i_b}} \right) + \frac{C^2}{2} \left(e^{i\frac{2\pi m}{i_b}} - 2 + e^{-i\frac{2\pi m}{i_b}} \right)}{1} \right| \leq 1. \quad (97)$$

Entonces, desarrollando, se obtiene

$$\begin{aligned} 1 &\geq \left| 1 - iC \sin\left(\frac{2\pi m}{i_b}\right) + C^2 \left(\cos\left(\frac{2\pi m}{i_b}\right) - 1 \right) \right|, \\ 1 &\geq \sqrt{\left(1 + C^2 \left(\cos\left(\frac{2\pi m}{i_b}\right) - 1 \right)\right)^2 + C^2 \sin^2\left(\frac{2\pi m}{i_b}\right)}, \\ 1 &\geq 1 + 2C^2 \cos\left(\frac{2\pi m}{i_b}\right) - 2C^2 + C^4 \cos^2\left(\frac{2\pi m}{i_b}\right) - 2C^4 \cos\left(\frac{2\pi m}{i_b}\right) + C^4 + C^2 \left(1 - \cos^2\left(\frac{2\pi m}{i_b}\right)\right), \\ 0 &\geq -(C^2 - C^4) + 2(C^2 - C^4) \cos\left(\frac{2\pi m}{i_b}\right) - (C^2 - C^4) \cos^2\left(\frac{2\pi m}{i_b}\right), \\ 0 &\leq (1 - C^2) \left(\cos\left(\frac{2\pi m}{i_b}\right) - 1 \right)^2. \end{aligned}$$

Ahora, como el segundo factor en la última desigualdad es siempre no negativo entonces el primer factor $(1 - C^2)$ debe ser también no negativo y eso significa que la condición de Von Neumann en este caso se reduce nuevamente a

$$0 \leq |C| \leq 1.$$

Dicha restricción determina entonces los casos en los cuales este esquema numérico es estable, y por lo tanto convergente. Del mismo modo que en los esquemas anteriores, la restricción impuesta por la prueba de Von Neumann coincide con la restricción de que la causalidad numérica no contradiga a la causalidad física, es decir, que la velocidad con la cual se propaga la información en el esquema numérico no supere a la velocidad con la cual se propaga la solución de D'Alembert.

Adicionalmente, como el método es explícito podemos también expresarlo mediante una suma del tipo expuesto en la ecuación (60) y los coeficientes resultan ser:

$$\begin{aligned} b_{-1} &= \frac{C}{2} (C + 1), \\ b_0 &= 1 - C^2, \\ b_1 &= \frac{C}{2} (C - 1). \end{aligned}$$

Como la única restricción para implementar este sistema es $0 \leq |C| \leq 1$, entonces los coeficientes b_k pueden ser negativos y por lo tanto el esquema no es monótono.

Por último, observamos en la expresión (94) que el esquema de Lax-Wendroff se reduce a

$$U_i^{j+1} = U_{i-1}^j \quad (98)$$

en el caso de que $C = 1$. Esto significa que los valores de los puntos simplemente se recorren un lugar hacia la

izquierda con cada paso temporal y, como se ha visto en la sección 5.1, implica también que en este caso la solución numérica U es idéntica a la solución exacta u , ecuación (84). Además, del mismo modo que los esquemas anteriores, la ecuación (95) correspondiente al error de truncamiento local del al esquema de Lax-Wendroff se reduce a

$$L = \frac{1}{\Delta t} (u(x, t + \Delta t) - u(x - c\Delta t, t))$$

cuando $C = 1$, y como se ha expuesto también en la sección 5.1 esto implica que $L = 0$, como podía esperarse.

5.4. Esquema de Wendroff.

Este esquema numérico es un poco más complicado que los esquemas anteriores debido a que no es *explícito* sino *implícito*. También tiene, sin embargo, una ventaja sobre los esquemas anteriores, y es que, como será mostrado al final de esta sección, su implementación no está sujeta a condición alguna sobre C . Para construir este esquema nuevamente recurrimos a la construcción de promedios; en este caso tomamos el promedio de la aproximación a la derivada espacial en dos diferentes niveles temporales, y el promedio de la derivada temporal en dos diferentes puntos espaciales. De este modo, al sustituir las aproximaciones en la ecuación de advección (17),

$$u_t + cu_x = 0,$$

se obtiene

$$\frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_i^{j+1} - U_i^j}{\Delta t} \right) + c \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right) = 0, \quad (99)$$

o equivalentemente,

$$(1 + C) U_{i+1}^{j+1} + (1 - C) U_i^{j+1} = (1 - C) U_{i+1}^j + (1 + C) U_i^j. \quad (100)$$

En este caso están involucrados en la expresión dos puntos del nivel temporal t_{j+1} por lo que para implementar el esquema podríamos despejar U_i^{j+1} o U_{i+1}^{j+1} . Como hemos dicho en la sección 5.1, podemos usar un criterio de causalidad, tomando en cuenta que sabemos de antemano que la información de la señal viaja de izquierda a derecha; es decir, despejamos U_{i+1}^{j+1} :

$$\begin{aligned} U_{i+1}^{j+1} &= \frac{1}{(1 + C)} \left((1 - C) U_{i+1}^j + (1 + C) U_i^j - (1 - C) U_i^{j+1} \right) \\ &= \frac{1 - C}{1 + C} (U_{i+1}^j - U_i^{j+1}) + U_i^j. \end{aligned} \quad (101)$$

Por causalidad en este caso nos referimos a la señal que describe la solución analítica de la ecuación de advección y no a la causalidad numérica que será mencionada abajo.

Esta expresión determina la relación causal entre puntos vecinos que gráficamente puede observarse en la figura 25. Así mismo en la figura 26 se encuentran los diagramas que nos permiten discriminar los puntos de la malla sobre los cuales no es necesario hacer cálculos.

Podemos observar en la figura 25 dos puntos en el nivel temporal t_{j+1} , lo cual implica que en este esquema numérico la condición inicial no es suficiente para calcular el valor de U en puntos de los niveles temporales superiores, por esta razón en los diagramas de la figura 26 se han colocado cuadros negros (ceros) en el extremo

izquierdo, una discusión al respecto de este punto puede encontrarse en la sección 5.5.3.

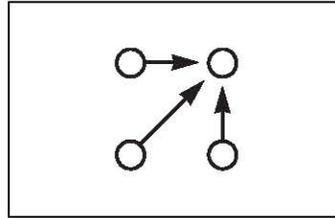


Figura 25: Relación causal numérica entre puntos vecinos para el esquema numérico de Wendroff.

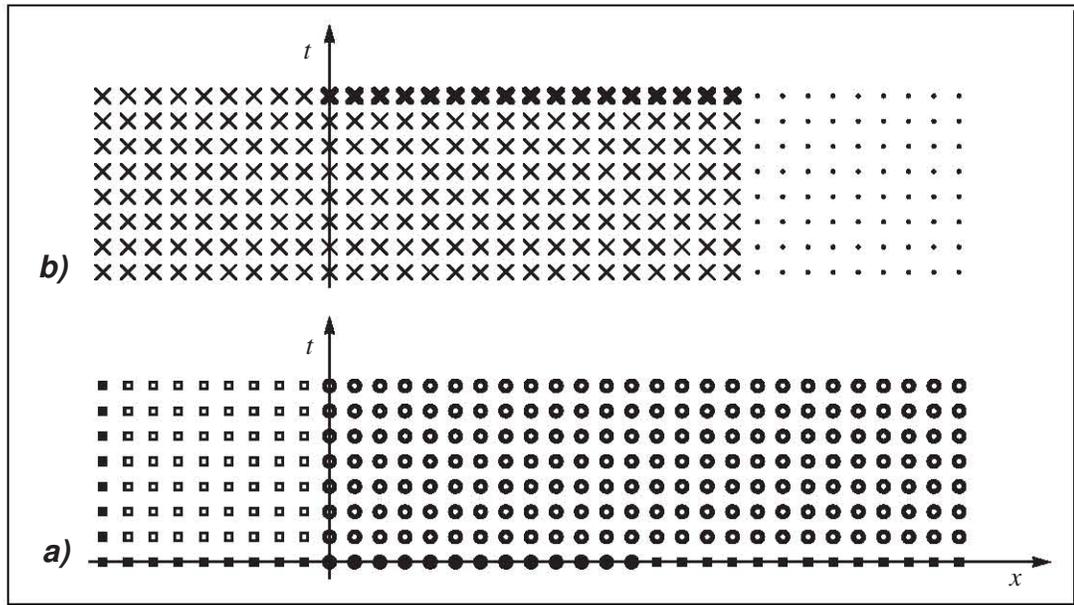


Figura 26: Diagramas de causalidad numérica para el esquema de Wendroff: a) primero y b) segundo. Para la definición de esta notación vea la sección 4.13 y la sección 5.5.3.

Para calcular el error de truncamiento local correspondiente a este esquema partimos de su expresión algebraica, que es

$$0 = \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_i^{j+1} - U_i^j}{\Delta t} \right) + c \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right).$$

Luego, como ya hemos visto antes, sustituimos la solución de la ecuación de advección evaluada en los puntos correspondientes,

$$L = \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x + \Delta x, t)}{\Delta t} + \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \right) + c \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x, t + \Delta t)}{\Delta x} + \frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} \right). \quad (102)$$

Como se trata de una expresión muy larga, expandimos por separado cada uno de los términos en su serie de Taylor:

$$\begin{aligned} \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x + \Delta x, t)}{\Delta t} \right) &= \frac{1}{2} u_t + \frac{1}{2} (\Delta x) u_{xt} + \frac{1}{4} (\Delta t) u_{tt} \\ &\quad + \frac{1}{12} \left(3 (\Delta x)^2 u_{xxt} + 3 (\Delta x) (\Delta t) u_{xtt} + (\Delta t)^2 u_{ttt} \right) \\ &\quad + O\left((\Delta x)^3 + (\Delta t)^3\right), \end{aligned} \quad (103)$$

$$\frac{1}{2} \left(\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \right) = \frac{1}{2} u_t + \frac{1}{4} (\Delta t) u_{tt} + \frac{1}{12} (\Delta t)^2 u_{ttt} + O\left((\Delta t)^3\right), \quad (104)$$

$$\begin{aligned} c \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x, t + \Delta t)}{\Delta x} \right) &= \frac{c}{2} u_x + \frac{c}{4} (\Delta x) u_{xx} + \frac{c}{2} (\Delta t) u_{xt} \\ &\quad + \frac{c}{12} \left((\Delta x)^2 u_{xxx} + 3 (\Delta x) (\Delta t) u_{xxt} + 3 (\Delta t)^2 u_{xtt} \right) \\ &\quad + O\left((\Delta x)^3 + (\Delta t)^3\right), \end{aligned} \quad (105)$$

$$c \frac{1}{2} \left(\frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} \right) = \frac{c}{2} u_x + \frac{c}{4} (\Delta x) u_{xx} + \frac{c}{12} (\Delta x)^2 u_{xxx} + O\left((\Delta x)^3\right). \quad (106)$$

El siguiente paso es aprovechar el hecho de que u satisface la ecuación de advección y por lo tanto debe satisfacer cada una de las relaciones (62). Esto nos permite expresar las ecuaciones (103), (104), (105) y (106) en términos únicamente de derivadas parciales con respecto a x :

$$\begin{aligned} \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x + \Delta x, t)}{\Delta t} \right) &= -\frac{c}{2} u_x - \frac{c}{2} (\Delta x) u_{xx} + \frac{c^2}{4} (\Delta t) u_{xx} \\ &\quad + \frac{1}{12} \left(-3c (\Delta x)^2 u_{xxx} + 3c^2 (\Delta x) (\Delta t) u_{xxx} + c^3 (\Delta t)^2 u_{xxx} \right) \\ &\quad + O\left((\Delta x)^3 + (\Delta t)^3\right), \end{aligned}$$

$$\frac{1}{2} \left(\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \right) = -\frac{c}{2} u_x + \frac{c^2}{4} (\Delta t) u_{xx} - \frac{c^3}{12} (\Delta t)^2 u_{xxx} + O\left((\Delta t)^3\right),$$

$$\begin{aligned} c \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x, t + \Delta t)}{\Delta x} \right) &= \frac{c}{2} u_x + \frac{c}{4} (\Delta x) u_{xx} - \frac{c^2}{2} (\Delta t) u_{xx} \\ &\quad + \frac{1}{12} \left(c (\Delta x)^2 u_{xxx} + 3c^2 (\Delta x) (\Delta t) u_{xxx} + 3c^3 (\Delta t)^2 u_{xxx} \right) \\ &\quad + O\left((\Delta x)^3 + (\Delta t)^3\right), \end{aligned}$$

$$c \frac{1}{2} \left(\frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} \right) = \frac{c}{2} u_x + \frac{c}{4} (\Delta x) u_{xx} + \frac{c}{12} (\Delta x)^2 u_{xxx} + O\left((\Delta x)^3\right).$$

Finalmente, sumando las cuatro expresiones anteriores se cancelan varios términos y obtenemos el error de

truncamiento local en la siguiente forma:

$$\begin{aligned} L = & u_{xxx} \left(\frac{1}{12} (-3c (\Delta x)^2 + 3c^2 (\Delta x) (\Delta t) - c^3 (\Delta t)^2) - \frac{c^3}{12} (\Delta t)^2 \right. \\ & + \frac{1}{12} (c (\Delta x)^2 - 3c^2 (\Delta x) (\Delta t) + 3c^3 (\Delta t)^2) + \frac{c}{12} (\Delta x)^2 \left. \right) \\ & + O((\Delta x)^3 + (\Delta t)^3), \end{aligned}$$

donde observamos que este esquema es consistente con la ecuación de advección puesto que cuando Δx y Δt tienden a cero L también tiende a cero.

Para conocer el orden de este esquema numérico expresamos L en términos de Δt bajo el supuesto de que el cociente $\frac{\Delta x}{\Delta t}$ es una constante. Agrupando términos queda:

$$L = O((\Delta t)^2).$$

Por lo tanto, al igual que el esquema de Lax-Wendroff, el esquema de Wendroff es de segundo orden.

A partir de la expresión (100) se observa que los polinomios que describen este esquema numérico son:

$$\begin{aligned} W &= (1 + C) S_+ + (1 - C), \\ V &= (1 - C) S_+ + (1 + C). \end{aligned}$$

O sea que (100) se escribe como

$$W(S_-, S_+) U_i^{j+1} = V(S_-, S_+) U_i^j.$$

La condición de estabilidad de Von Neumann en este caso adquiere la forma

$$\left| \frac{(1 - C) e^{i \frac{2\pi m}{b}} + (1 + C)}{(1 + C) e^{i \frac{2\pi m}{b}} + (1 - C)} \right| \leq 1. \quad (107)$$

Para simplificar el cálculo recordamos tres propiedades de los números complejos: primero que la norma de un producto es el producto de las normas; segundo, que la norma de cualquier número es igual a la norma de su conjugado; y tercero, que la norma de un número que pueda escribirse como $e^{i\theta}$, $\theta \in \mathbb{R}$ es uno.

Entonces, desarrollando,

$$\left| (1 - C) e^{i \frac{2\pi m}{b}} + (1 + C) \right| = \left| \left((1 - C) e^{i \frac{2\pi m}{b}} + (1 + C) \right) e^{-i \frac{2\pi m}{b}} \right| \quad (108)$$

$$= \left| (1 - C) + (1 + C) e^{-i \frac{2\pi m}{b}} \right| \quad (109)$$

$$= \left| (1 - C) + (1 + C) e^{i \frac{2\pi m}{b}} \right|, \quad (110)$$

y por lo tanto

$$\left| \frac{(1 - C) e^{i \frac{2\pi m}{b}} + (1 + C)}{(1 + C) e^{i \frac{2\pi m}{b}} + (1 - C)} \right| = 1,$$

independientemente del valor de C . Es decir, a diferencia de los tres esquemas anteriores, para el esquema de Wendroff la condición de Von Neumann se satisface incondicionalmente y por lo tanto este esquema es incondicionalmente estable.

Del mismo modo, para este esquema numérico el valor de C no determina si la causalidad numérica contradice a la causalidad física. Eso se debe a que su carácter implícito, como será mostrado en la sección 5.5.3, implica que para obtener el valor de U en un determinado punto se emplea el valor de muchos otros puntos del nivel temporal anterior (vea también la sección 4.7).

Al igual que los esquemas anteriores el esquema de Wendroff se simplifica bastante en el caso $C = 1$, para el cual la ecuación (101) toma la forma

$$U_{i+1}^{j+1} = U_i^j ; \quad (111)$$

es decir, cuando se avanza un nivel temporal los valores de los puntos se recorren un lugar hacia la derecha, y de la misma manera que se ha expuesto en la sección 5.1, esto implica que para el caso $C = 1$ la solución numérica de este esquema es precisamente la solución exacta u , ecuación (84). Para este caso también es posible demostrar que se tiene $L = 0$, para ello se impone $C = 1$ en la ecuación (102), y después de una traslación de la variable t se tiene

$$L = \frac{1}{\Delta t} (u(x, t + \Delta t) - u(x - c\Delta t, t)) ,$$

que como se ha mostrado en la sección 5.1 implica precisamente $L = 0$.

5.5. Implementación.

Además de las precisiones que se han hecho al respecto de los esquemas numéricos en las secciones anteriores, al momento de escribir un código, para que una computadora pueda hacer la propagación numérica es necesario tomar algunas decisiones que serán detalladas a lo largo de esta sección.

5.5.1. Parámetros de control, fijos y de evaluación.

La solución exacta u de la ecuación de advección es una función de las variables x, t . Del mismo modo, U como aproximación de u es una función de las mismas variables, pero de modo equivalente, tal como se ha manejado en las secciones anteriores, es posible expresarlo por medio de una notación de super y subíndices i, j

$$U_i^j$$

en lugar de escribir la dependencia explícita, como

$$U(x_L + (i - 1)\Delta x, (j - 1)\Delta t)$$

porque existe una relación directa (y lineal) entre x e i y entre t y j :

$$x = x_L + (i - 1)\Delta x$$

$$t = (j - 1)\Delta t$$

cuya descripción puede encontrarse en la sección 4.2.

Para los esquemas numéricos relativos a la ecuación de advección uno de los aspectos que resulta conveniente cuando se expresa U como función de los índices i, j es que después de haber calculado U_i^j para un esquema numérico usando ciertos valores específicos $\Delta x, \Delta t, c$, digamos

$$\begin{aligned}\Delta x &= \Delta x_0, \\ \Delta t &= \Delta t_0, \\ c &= c_0\end{aligned}$$

puede conocerse el valor de U_i^j que se obtendría a partir de valores diferentes $\Delta x_1, \Delta t_1, c_1$ siempre que la variable $C = \frac{c\Delta t}{\Delta x}$ tenga el mismo valor, en otras palabras, siempre que:

$$\frac{c_1\Delta t_1}{\Delta x_1} = \frac{c_0\Delta t_0}{\Delta x_0}$$

y esto se debe a que, en los esquemas numéricos que se han descrito en secciones anteriores, el valor de U_i^j depende solamente de C y de los valores de U en los puntos vecinos, como puede observarse en las ecuaciones (78), (87), (94) y (101). Sin embargo, es importante notar que en el espacio (x, t) el punto al cual corresponde un determinado valor U_i^j efectivamente depende de los valores específicos de Δx y Δt , por ejemplo en el primer caso dicho punto correspondería a

$$\begin{aligned}x &= x_L + (i - 1) \Delta x_0, \\ t &= (j - 1) \Delta t_0,\end{aligned}$$

y en el segundo caso correspondería a

$$\begin{aligned}x &= x_L + (i - 1) \Delta x_1, \\ t &= (j - 1) \Delta t_1.\end{aligned}$$

Lo importante de estas observaciones es que permiten restringir el campo de estudio, es decir, para poder observar los diferentes resultados que se obtienen de la implementación de los esquemas numéricos, no hace falta usar diferentes valores de $\Delta x, \Delta t$ y c , sino simplemente diferentes valores de C y tomar en cuenta que para recuperar la dependencia al respecto de las variables originales x, t haría falta especificar los valores de $\Delta x, \Delta t$ y c .

Otra observación que nos permite restringir el campo de estudio en este caso, es que debido al carácter lineal de los esquemas numéricos relativos a la ecuación de advección, expuesto en la sección 4.12.3 mediante la ecuación (74), no es relevante la amplitud de la condición inicial y por esta razón no se incluye este valor como parámetro de control.

Por otro lado podemos esperar que mientras más puntos hayamos usado para describir la condición inicial los resultados obtenidos serán mejores. Variaremos entonces el valor de i_p , que para condiciones iniciales periódicas se define como el número de puntos espaciales que hay en un periodo (ver figura 12), y para condiciones iniciales de soporte compacto es el número de puntos espaciales que hay a lo largo del soporte (ver figura 13). Estas definiciones nos permiten comparar los resultados obtenidos para los dos tipos de condiciones iniciales.

Del mismo modo podemos verificar que para ambos tipos de condiciones iniciales se cumple

$$\Delta x = \frac{1}{i_p},$$

por lo que la elección de un determinado valor de i_p determina también el valor de Δx .

Además, debido a que

$$C = c \frac{\Delta t}{\Delta x},$$

al escoger un valor para C queda igualmente determinado el valor de Δt puesto que hemos dispuesto, por simplicidad, que c tenga el valor de 1 en todos los casos.

Entonces nuestros *parámetros de control* son dos, i_p y C . En el caso de i_p se han escogido 6 valores en una escala logarítmica entre $i_p = 10$ y $i_p = 1000$, que son: 10, 25, 63, 158, 400, 1000 (cada número es aproximadamente 2.5 veces el anterior).

En el caso de C se han empleado dos criterios: primero, buscando valores en una escala logarítmica entre $C = 0.01$ y $C = 5$; y segundo, buscando valores cerca de $C = 1$, que como hemos visto, implica comportamientos especiales para los diferentes esquemas numéricos. Los valores que se han escogido son: 0.01, 0.1, 0.5, 1, 1.1, 5, por lo que tendremos 36 combinaciones posibles de valores de C e i_p para cada tipo de condición inicial.

Las cantidades que nos permitirán comparar los resultados numéricos entre sí, a las cuales en adelante llamaremos *parámetros de evaluación*, son las siguientes:

Amplitud relativa: Se mide como la diferencia entre el máximo y el mínimo de la solución numérica U en un determinado nivel temporal t_{j_0} sobre la amplitud de la condición inicial.

$$A = \frac{\max \{U_i^j | j = j_0\} - \min \{U_i^j | j = j_0\}}{\text{Amplitud de la condición inicial}}$$

Desplazamiento: Es el corrimiento del máximo de la solución numérica U con respecto al corrimiento del máximo de la solución analítica u , para un determinado nivel temporal:

$$D = i_{\max} \Delta x - x_{\max},$$

donde i_{\max} es el valor de i que corresponde al máximo de U y x_{\max} es el valor de x que corresponde al máximo de u en dicho nivel temporal.

Norma: Se define como la máxima diferencia entre la solución numérica U y la solución exacta u , para un determinado nivel temporal:

$$N = \max \left\{ \left| U_i^j - u(i\Delta x, j\Delta t) \right| \mid j = j_0 \right\}.$$

Finalmente señalamos los *parámetros fijos* que tienen el mismo valor para todas las propagaciones numéricas:

Constante c : Es la constante que aparece en la ecuación de advección (17), esta constante determina la velocidad de propagación de la solución exacta u . El valor que se le ha dado es $c = 1$, escogido por simplicidad y sin pérdida de generalidad por la discusión mencionada con respecto al parámetro C .

Pasos temporales calculados: Es la cantidad de pasos temporales que efectivamente se han obtenido mediante cálculo numérico, es decir, el máximo valor de j . Este valor se ha escogido en 2500, porque un número mucho mayor hubiera sido muy difícil de manejar en términos computacionales, y un número mucho menor hubiera disminuido significativamente la posibilidad de observar diferencias en los resultados de los diferentes esquemas numéricos.

Se ha mantenido fijo también el *ancho de la condición inicial*; es decir, para el caso de condiciones iniciales periódicas la longitud del periodo, y para el caso de condiciones iniciales de soporte compacto la longitud del soporte mismo. Como puede observarse en las figuras 12 y 13 este parámetro en ambos casos ha sido tomado igual a 1.

5.5.2. Implementación de los esquemas explícitos: contracorriente, Lax-Friedrichs y Lax-Wendroff.

Para calcular el valor de U en un dominio de propagación a partir de la condición inicial según los esquemas contracorriente, Lax-Friedrichs y Lax-Wendroff se usan las expresiones (78), (87) y (94) respectivamente.

$$\begin{aligned} \text{Contraorriente:} \quad U_i^{j+1} &= U_i^j - C(U_i^j - U_{i-1}^j) \\ \text{Lax-Friedrichs:} \quad U_i^{j+1} &= U_{i-1}^j (1 + C) \frac{1}{2} + U_{i+1}^j (1 - C) \frac{1}{2} \\ \text{Lax-Wendroff:} \quad U_i^{j+1} &= U_i^j - \frac{C}{2}(U_{i+1}^j - U_{i-1}^j) + \frac{C^2}{2}(U_{i+1}^j - 2U_i^j + U_{i-1}^j) \end{aligned}$$

Se toman los valores de la condición inicial en primera instancia para calcular los correspondientes al segundo nivel temporal, luego se aplica el mismo procedimiento sobre estos valores que se han calculado para obtener el tercer nivel temporal y así sucesivamente.

Aunque el procedimiento es sencillo, es necesario tener cuidado al llegar a los puntos que se encuentran en los extremos. Como se ha mencionado, el tratamiento de estos puntos depende del tipo de condición inicial que se esté usando.

Si la condición inicial es periódica se usa el valor del extremo opuesto; por ejemplo, en el caso de Lax-Friedrichs, para calcular el valor de U_1^{j+1} se usan los valores de $U_{i_p}^j$ y U_2^j como se muestra en la siguiente expresión:

$$U_1^{j+1} = U_{i_p}^j (1 + C) \frac{1}{2} + U_2^j (1 - C) \frac{1}{2}.$$

Para calcular el valor de $U_{i_p}^{j+1}$ se usan los valores de $U_{i_p-1}^j$ y U_1^j de la siguiente forma:

$$U_{i_p}^{j+1} = U_{i_p-1}^j (1 + C) \frac{1}{2} + U_1^j (1 - C) \frac{1}{2}.$$

Por otro lado si la condición inicial es de soporte compacto es necesario tomar en cuenta que solamente se calcularán los valores de U correspondientes a la región donde intersectan las cruces (X o bien X) y las circunferencias ● de de los diagramas de causalidad (figuras 20, 22 y 24). Algunos de los puntos que se encuentran en la orilla de esta región dependen de los valores de puntos que están fuera de la misma pero, por el modo como hemos construido los diagramas, el valor de U correspondiente a estos puntos externos será siempre cero, de modo que, para

hacer los cálculos de los puntos al extremo simplemente, cuando sea necesario, se colocan ceros en el lugar que les corresponde.

5.5.3. Implementación del esquema implícito: Wendroff.

Para implementar el esquema de Wendroff usamos la ecuación (101)

$$U_{i+1}^{j+1} = \frac{1-C}{1+C} (U_{i+1}^j - U_i^{j+1}) + U_i^j,$$

donde se parte también de una condición inicial, pero en este caso eso no es suficiente porque del lado derecho de la ecuación siempre hará falta el valor de U_i^{j+1} . Aunque este es un problema que se presenta en todos los puntos del nivel temporal t_{j+1} , sería suficiente conocer U_1^{j+1} porque con este valor podría obtenerse U_2^{j+1} y luego así el resto de los puntos del nivel t_{j+1} . Entonces el problema se reduce a encontrar el valor de U_1^{j+1} , lo obtendremos de maneras diferentes dependiendo del tipo de condiciones iniciales que se tengan.

Para el caso de condiciones iniciales de soporte compacto obtendremos U_1^{j+1} suponiendo que muy lejos del pulso (soporte de la condición inicial), U vale cero para todos los niveles temporales. De cierta forma, ésta es una condición de frontera al infinito. Esta suposición resulta útil porque significa que podemos colocar una columna de puntos tales que $U = 0$, suficientemente lejos, en particular en el extremo izquierdo de la malla. Abusando ligeramente de la notación hemos usado en la figura 26 también cuadros negros ■ para denotar estos puntos.

Si ahora usamos la expresión (101) sobre los puntos del segundo nivel temporal a partir de los cuadros negros recién colocados obtendremos ceros hasta llegar a $x = 0$, y se repite el procedimiento para el resto de los niveles temporales, por lo que se han designado en la figura 26 con cuadros blancos □ todos estos puntos. Por lo tanto se puede obtener el valor de U_1^{j+1} simplemente considerando nulos los valores correspondientes a los puntos que están a la izquierda, es decir

$$U_1^{j+1} = \frac{1-C}{1+C} (U_1^j - 0) + 0. \quad (112)$$

Para condiciones iniciales periódicas como se ha hecho con los esquemas numéricos anteriores se usan los puntos de los extremos opuestos, que en este caso solamente se necesita para el punto del extremo izquierdo,

$$\begin{aligned} U_1^{j+1} &= \frac{1-C}{1+C} (U_1^j - U_{i_p}^{j+1}) + U_{i_p}^j \\ &= \frac{1-C}{1+C} U_1^j + U_{i_p}^j - \frac{1-C}{1+C} U_{i_p}^{j+1}, \end{aligned}$$

y para resolver el problema arriba mencionado, en este caso al respecto de $U_{i_p}^{j+1}$, se usa la misma expresión (101) para dejar U_1^{j+1} en términos de $U_{i_p-1}^{j+1}$

$$U_1^{j+1} = \frac{1-C}{1+C} U_1^j + U_{i_p}^j - \frac{1-C}{1+C} \left(\frac{1-C}{1+C} U_{i_p}^j + U_{i_p-1}^j - \frac{1-C}{1+C} U_{i_p-1}^{j+1} \right).$$

Si se realiza de manera iterativa una sustitución de este tipo, eventualmente se tendrá una expresión donde del lado derecho aparece nuevamente U_1^{j+1} . Llegado este momento, U_1^{j+1} será el único valor de U correspondiente al

nivel temporal t_{j+1} dentro de la ecuación, y por lo tanto podemos despejarlo para obtener

$$U_1^{j+1} = \frac{\left(\frac{1-C}{1+C}\right) U_1^j + U_{i_p}^j + \sum_{i=1}^{i_p-1} (-1)^i \left(\left(\frac{1-C}{1+C}\right)^{i+1} U_{i_p+1-i}^j + \left(\frac{1-C}{1+C}\right)^i U_{i_p-i}^j \right)}{1 + (-1)^{i_p+1} \left(\frac{1-C}{1+C}\right)^{i_p}}. \quad (113)$$

Como se ha mencionado, para ambos tipos de condiciones iniciales, una vez obtenido U_1^{j+1} es posible calcular el resto de los valores y así implementar este esquema numérico.

A continuación se presentan otros dos procedimientos de implementación que teóricamente son correctos, y desde un punto de vista práctico resultan de menor o igual rendimiento que el procedimiento arriba descrito. Una breve descripción de ambos se incluye para hacer más completa la discusión de métodos numéricos implícitos.

Primero, para el caso de condiciones iniciales periódicas se hace una sustitución recursiva en orden descendente, $i_p, i_p - 1, i_p - 2, i_p - 3, \dots, 1$. Ahora, bien podríamos haberlo hecho en orden ascendente en lugar de descendente. Lo que resulta inconveniente del caso ascendente es que la expresión final para U_1^{j+1} , análoga a la expresión (113), involucra factores del tipo

$$\left(\frac{1+C}{1-C}\right)^{i_p},$$

que, además de presentar problemas cuando $C = 1$, en la mayoría de los casos resultan ser demasiado grandes porque $c > 0$ implica que

$$\frac{1+C}{1-C} > 1,$$

e i_p llega incluso a valer 10^3 . De este modo fácilmente puede alcanzarse el límite superior para el valor de una variable numérica (cerca de 10^{300} en el caso de MATLAB 7). Por lo tanto no es conveniente implementar el esquema numérico de esta manera.

Por otro lado, esta observación nos lleva a optimizar la implementación de la ecuación (113) debido a que también existe un límite inferior para el valor absoluto de las variables numéricas cuyo valor es distinto de cero. En otras palabras, para MATLAB 7 no es posible diferenciar entre cero y 10^{-350} ; por lo tanto, en la suma que aparece dentro de la ecuación (113) es posible despreciar los términos que sean suficientemente pequeños, la cantidad de términos despreciados depende del valor de C .

La segunda alternativa se refiere al procedimiento que se ha descrito en la sección 4.7. Consistiría en expresar el esquema de Wendroff en notación matricial

$$\mathbf{A}U^{j+1} = \mathbf{B}U^j,$$

para luego obtener un esquema explícito equivalente por medio del cálculo de la inversa de \mathbf{A} . Lo inconveniente en este caso es que su implementación requiere muchas más operaciones que el procedimiento descrito al principio de esta sección. Primero, la inversión de la matriz \mathbf{A} por medio del método de LU descrito en el texto de Press [28] requiere aproximadamente $(i_b)^3$ operaciones, al decir *operación* nos referimos una suma y un producto, donde alguno de los factores puede estar elevado a una potencia, esta definición no es muy precisa pero funciona muy bien para hacer las estimaciones. La variable i_b es el número de puntos en el eje espacial y eso significa que las matrices \mathbf{A} y \mathbf{B} son $i_b \times i_b$. Luego, se requieren $3(i_b)^2$ operaciones para efectuar el producto $\mathbf{A}^{-1}\mathbf{B}$, considerando que \mathbf{B} es prácticamente tridiagonal (discutido en la sección 4.7), y finalmente se requieren $(i_b)^2$ operaciones para calcular el

producto de la matriz resultante con \mathbf{U}^j . Entonces, se requieren del orden de $(i_b)^3$ operaciones para iniciar el proceso y luego $(i_b)^2$ operaciones para cada paso temporal.

En el mismo texto de Press [28] sección 2.4 se sugiere un algoritmo, para el caso de que tanto \mathbf{A} como \mathbf{B} sean tridiagonales, que usa aproximadamente $7i_b$ operaciones para obtener \mathbf{U}^{j+1} sin obtener explícitamente \mathbf{A}^{-1} . Para implementar este algoritmo, además de que se debe hacer un pequeño tratamiento para que las matrices involucradas sean estrictamente tridiagonales, se requiere cierto cuidado al respecto de la relación entre los valores de la diagonal, de otro modo el algoritmo falla.

Por otro lado, en la primera implementación descrita de este esquema numérico se requieren en cada paso temporal cuando mucho $2i_b$ operaciones para encontrar U_1^j por medio de la expresión (113) (si se tienen condiciones iniciales de soporte compacto se requiere una sola operación para obtener U_1^j de acuerdo a la ecuación (112)) y luego $3i_b$ operaciones para calcular los valores del resto de los puntos por medio de la ecuación (101).

Lo que marca la diferencia real, con respecto al tiempo de computo cuando se efectúa la implementación, es el orden de magnitud de la cantidad de operaciones necesarias para cada paso temporal. Cuando se procede por medio de inversión de matrices por el método LU, cada paso temporal requiere del orden de $(i_b)^2$ operaciones; mientras que con el procedimiento que efectivamente se ha usado, o en el algoritmo sugerido en el texto de Press, solamente se requieren del orden de i_b operaciones (no debe olvidarse que i_b alcanza el valor de 10^3). Por lo tanto, se requeriría mucho más tiempo para hacer la implementación por medio de inversión directa de matrices.

5.6. Resultados.

Una vez implementados cada uno de los esquemas descritos, se obtienen conjuntos de valores para U asociados a puntos en una malla discreta, y estos los podemos graficar para niveles temporales dados junto a la solución de D'Alembert, que es una solución analítica exacta, como ya se ha visto, para evaluar su exactitud de acuerdo a la proximidad con esta solución exacta. Este tipo de gráficas se encuentran en la sección 5.6.1 y nos permiten comparar cualitativamente los resultados numéricos.

Para hacer un análisis detallado de los resultados es necesario usar los parámetros de evaluación descritos en la sección 5.5.1: la amplitud relativa A , el desplazamiento D , y la norma N . Dos de estos parámetros están relacionados a los conceptos de difusión y dispersión expuestos en la sección 2.4, y son la amplitud relativa A y el desplazamiento D , respectivamente. Como se ha visto en la sección 4.14 en este caso la presencia de dispersión y difusión en U necesariamente debe ser de origen numérico, es decir, dispersión numérica y difusión numérica, y su efecto puede estudiarse sobre propagaciones del tipo senoidal. Por otro lado, en la sección 4.12.3 se ha mostrado que las propagaciones basadas en condiciones iniciales de soporte compacto equivalen a sumas de propagaciones senoidales. Las propagaciones basadas en condiciones iniciales periódicas descritas en la sección 4.12.1 tienen precisamente esta forma senoidal, por lo que determinaremos la presencia de dispersión numérica y difusión numérica observando sus valores de amplitud relativa A y desplazamiento D . Luego, con esta información, buscaremos explicar las observaciones hechas en las propagaciones donde se tienen condiciones iniciales de soporte compacto.

En las gráficas presentadas de la sección 5.6.2 a la sección 5.6.5 usaremos como variable dependiente cada uno de los parámetros de evaluación A , N y D . Para el eje de las abscisas resultará más útil usar j como variable independiente, aunque podrían usarse también C o i_p .

Si se toma j como variable independiente (eje de las abscisas) para cada esquema numérico y cada parámetro

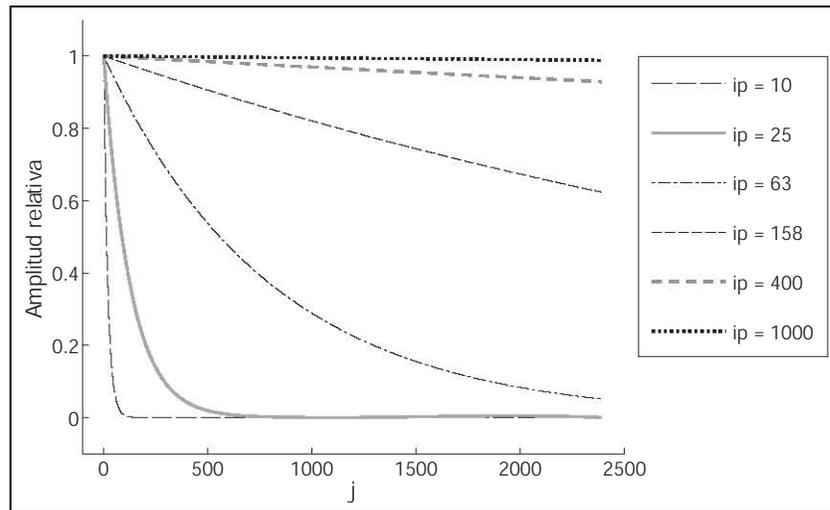


Figura 27: Amplitud relativa como función de j , para condiciones iniciales periódicas bajo el esquema contracorriente, con $C = 0.5$ y diferentes valores de i_p .

de evaluación tenemos 36 combinaciones de i_p y C para los casos de condiciones iniciales periódicas (y otras 36 para los casos de soporte compacto aunque por las razones expuestas antes observaremos primero los casos de condiciones iniciales periódicas). Si escogemos un esquema numérico y un valor para C nos quedan únicamente 6 gráficas posibles correspondientes a los 6 valores disponibles de i_p ; estas seis gráficas las podemos poner en una sola para obtener lo que se muestra en la figura 27. Lo que observaremos en las figuras de este tipo es el orden en el que aparecen las gráficas correspondientes a los diferentes valores de i_p con respecto al comportamiento ideal; son mejores los resultados que sean más próximos a dicho comportamiento. El comportamiento ideal es aquel que se observaría de haber tomado las medidas sobre la solución de D'Alembert; en el caso de la amplitud relativa A este sería un valor constante 1, y tanto en el caso del desplazamiento D como en el de la norma N el comportamiento ideal sería un valor constante igual a 0. Entonces, por ejemplo en la figura 27, se encuentra el mejor resultado en la gráfica correspondiente a $i_p = 1000$, en segundo lugar la correspondiente a $i_p = 400$, etc.

El valor de C que nos conduce a resultados más sencillos es $C = 1$. Para este caso se obtienen solamente gráficas que son líneas horizontales, donde ni la amplitud relativa, ni el desplazamiento, ni la norma cambian a medida que avanza j (se observa precisamente el comportamiento ideal), lo cual es consistente con las ecuaciones (82), (91), (98) y (111) que nos muestran que este valor de C reduce los esquemas numéricos a un corrimiento horizontal de los valores de U entre los puntos de la malla, donde además $U = u$.

Por otro lado, para los esquemas contracorriente, Lax-Friedrichs y Lax-Wendroff, los casos $C = 1.1$ y $C = 5$ tienen una interpretación sencilla: dichos esquemas en estas condiciones efectivamente son inestables, es decir, la amplitud de la señal crece rápidamente y sin cota superior. En la figura 28 puede observarse un ejemplo correspondiente a $C = 1.1$, $i_p = 158$. El hecho de que la amplitud crezca indefinidamente implica directamente que la norma N crece del mismo modo. De hecho, una vez que la amplitud de la señal comienza a crecer, la señal se deforma completamente, y por lo tanto tampoco tienen sentido las mediciones al respecto del desplazamiento D , por lo que en todos los análisis posteriores, a menos de que se indique lo contrario, cuando se hable de estos tres esquemas se elegirá $C \leq 1$. Es notable que incluso en estos casos inestables existe una región donde los resultados numéricos aproximan muy bien a la solución de D'Alembert. Dicha región está acotada por un valor de j (en cada caso diferente), después del cual se observa un crecimiento aproximadamente exponencial; es decir, se observan líneas rectas

cuando se usa un eje logarítmico para A .

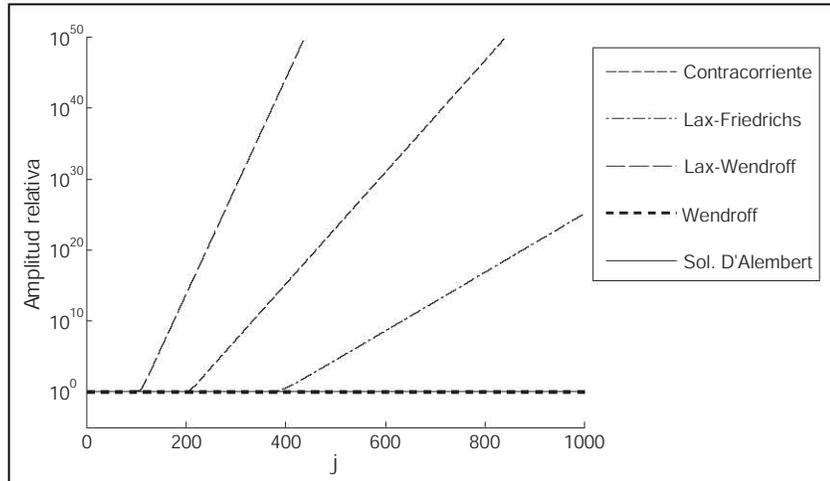


Figura 28: Amplitud relativa como función de j para un caso inestable con $C = 1.1$, donde $i_p = 158$.

La figura 28 no es del mismo tipo que la figura 27: en esta figura 28 se han fijado los valores de i_p y C para empalmar las gráficas correspondientes a los diferentes esquemas numéricos. Este segundo tipo de figura también será empleado en las secciones posteriores.

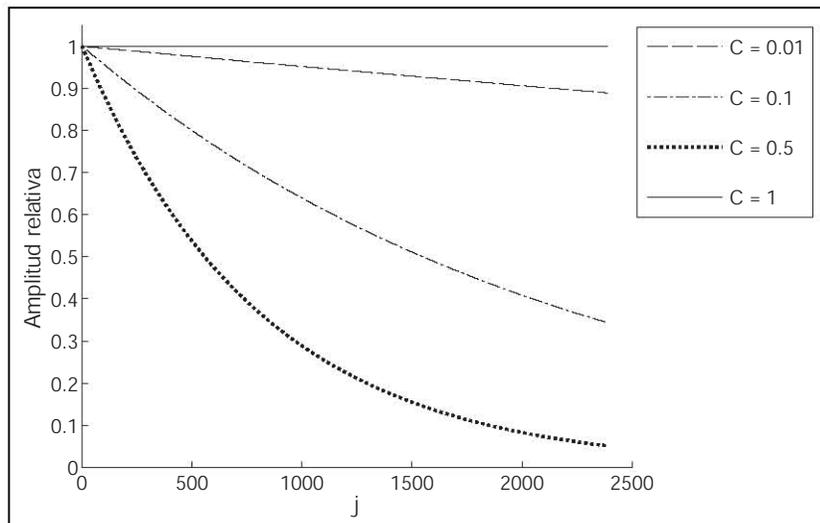


Figura 29: Amplitud relativa como función de j , para condiciones iniciales periódicas bajo el esquema contracorriente, con $i_p = 63$ y diferentes valores de C .

Entonces, nos quedan tres valores de C , que son 0.01, 0.1 y 0.5. Llevaremos a cabo figuras análogas a 27 para cada uno de estos valores de C y para cada uno de los distintos esquemas numéricos. Una tercera manera de elaborar las figuras es escoger un valor de i_p y empalmar las gráficas correspondientes a los tres valores de C que nos quedan, como se muestra en la figura 29. Nuevamente la idea es observar el orden en que aparecen las gráficas para determinar qué valor de C nos lleva a mejores resultados, es decir, resultados más próximos al comportamiento ideal. En la mencionada figura 29 la gráfica más alejada es la correspondiente a $C = 0.5$, la correspondiente a $C = 0.1$ es mejor, aún mejor es el caso $C = 0.01$ y, como se ha mencionado antes, el $C = 1$ es precisamente el caso ideal. La falta de monotonía que se presenta entre el orden de aparición de las gráficas y el valor de C será discutida más

adelante.

Adicionalmente a las gráficas expuestas en esta sección se proporciona material digital interactivo en un CD incluido con el presente texto que permite comparar una gama más amplia de resultados. Para una descripción de este material revise el apéndice A.

5.6.1. Gráficas de propagación.

Todas las gráficas presentadas en esta sección muestran en el eje de las abscisas la variable espacial x y en el eje de las ordenadas a los resultados numéricos U . Este tipo de gráficas permiten comparar directamente la solución de D'Alembert con los resultados numéricos para los diferentes esquemas numéricos con diferentes juegos de valores j, C, i_p .

La figuras que se presentan están organizadas en tres series de tres figuras cada una, en cada una de estas series están fijos dos de los parámetros mencionados (j, C, i_p) y varía el valor del tercero como se detalla a continuación. En las figuras 30, 31 y 32 se observan los perfiles de propagación obtenidos usando $C = 0.5$ e $i_p = 63$, para tres valores de j : 241, 721 y 2389. En esta primera serie se observa cómo a medida que transcurre el tiempo (a medida que aumenta el valor de j) se van separando cada vez más los resultados numéricos de la solución de D'Alembert. Para los esquemas contracorriente y Lax-Friedrichs es muy pronunciado el decrecimiento de la amplitud incluso desde el caso $j = 241$, mientras que para los otros dos esquemas dicho decrecimiento apenas es apreciable. Las figuras 33, 34 y 35 constituyen una segunda serie; en estos tres casos son constantes los valores de $i_p = 63$ y $j = 721$, pero cada uno corresponde a un valor diferente de C : 0.5, 0.1, 0.01. En esta serie se observa, exceptuando el caso del esquema de Lax-Friedrichs, que la proximidad mejora para valores de C más pequeños. Del mismo modo, las figuras 36, 37 y 38 constituyen la tercera serie; en todos estos casos se han usado los valores $C = 0.5$ y $j = 721$, variando en este caso i_p sobre los valores: 10, 63 y 400. Como podía esperarse, la proximidad es mejor en todos los casos para valores mayores de i_p .

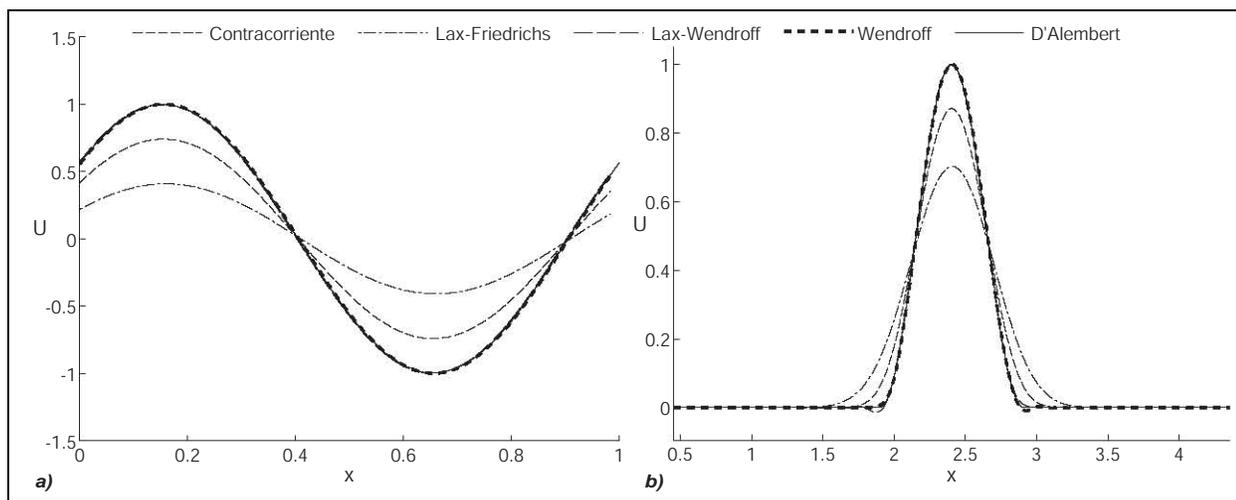


Figura 30: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 63$, $j = 241$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

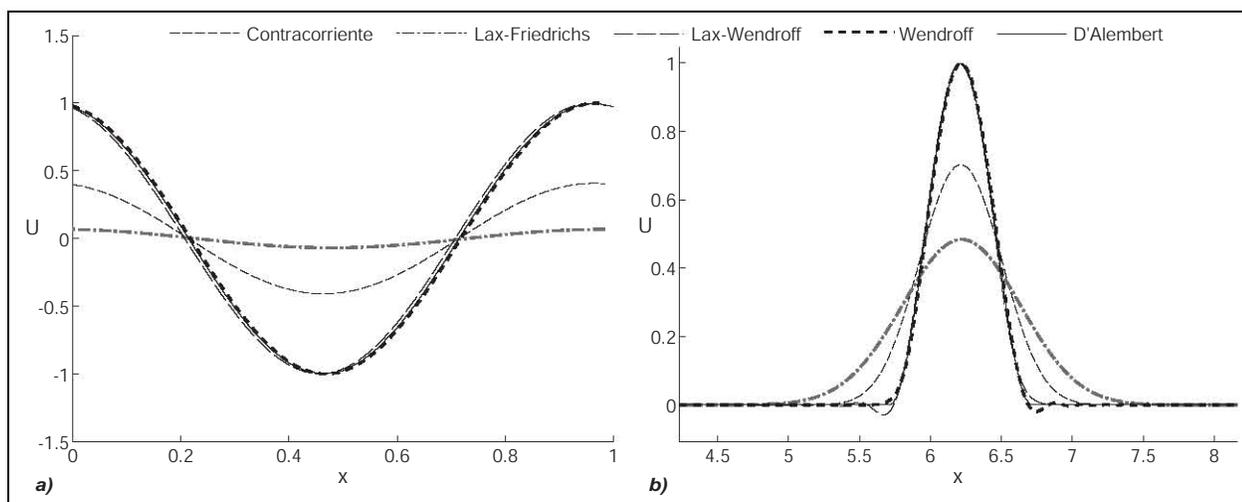


Figura 31: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 63$, $j = 721$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

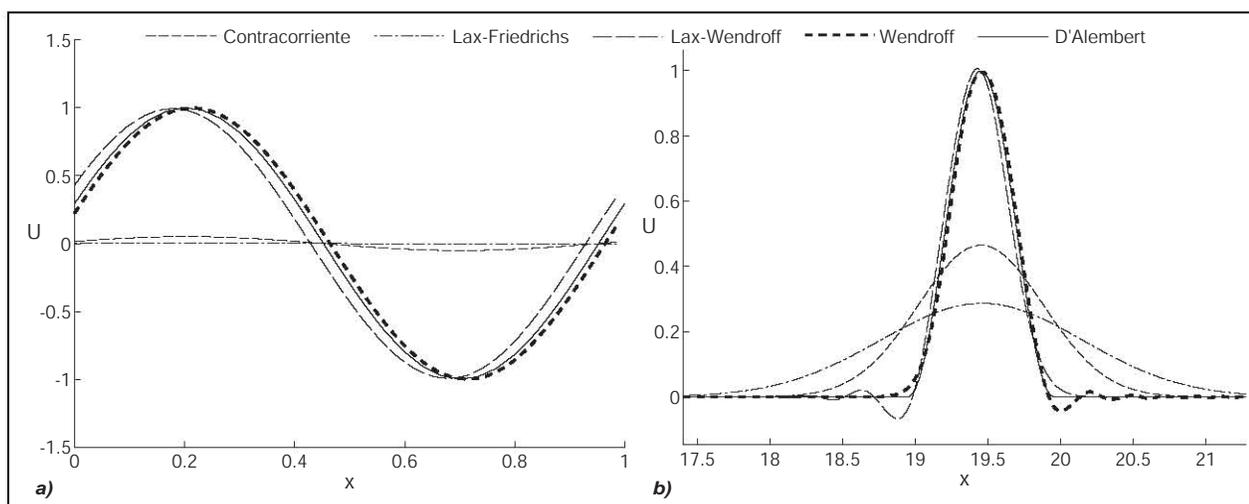


Figura 32: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 63$, $j = 2389$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

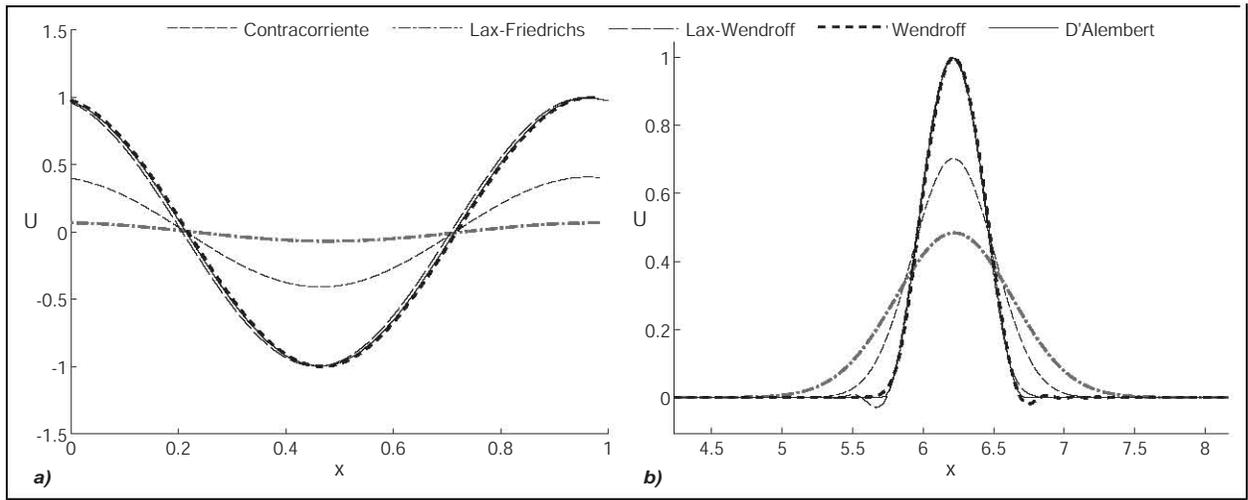


Figura 33: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 63$, $j = 721$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

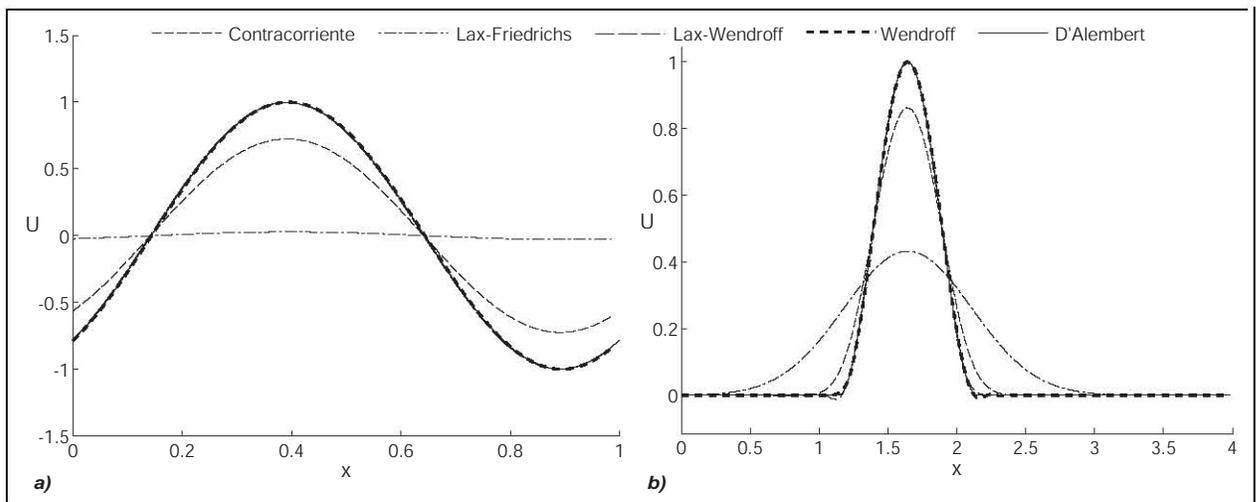


Figura 34: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.1$, $i_p = 63$, $j = 721$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

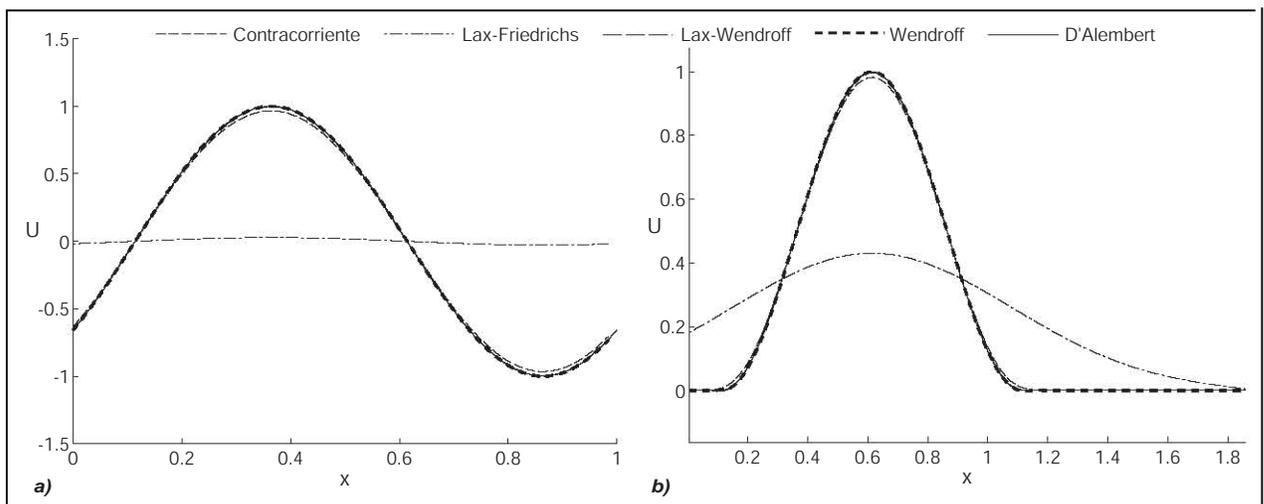


Figura 35: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.01$, $i_p = 63$, $j = 721$ respectivas a condiciones iniciales: **a)** periódicas, y **b)** de soporte compacto.

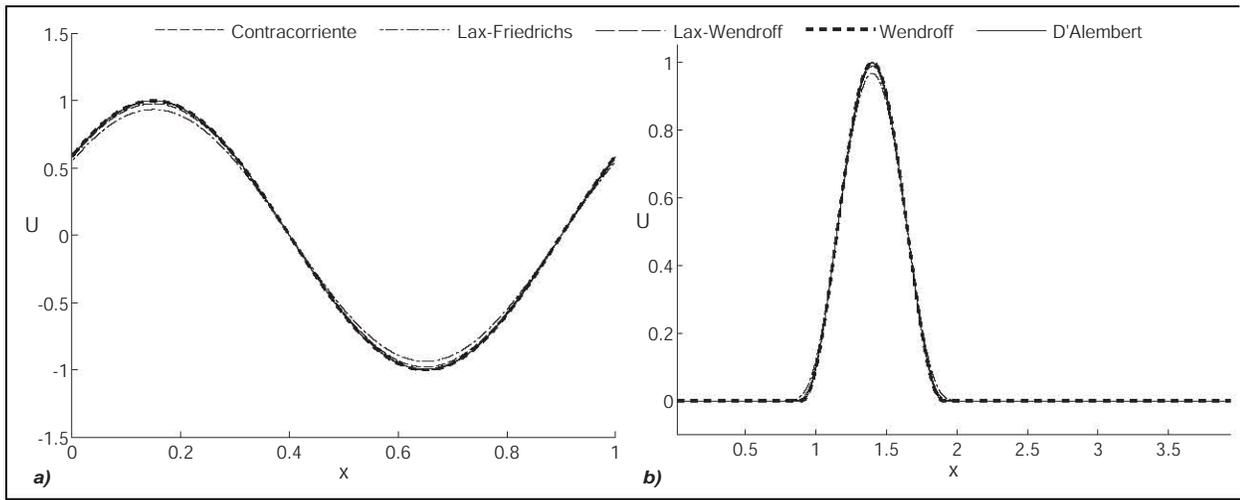


Figura 36: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 400$, $j = 721$ respectivas a condiciones iniciales: a) periódicas, y b) de soporte compacto.

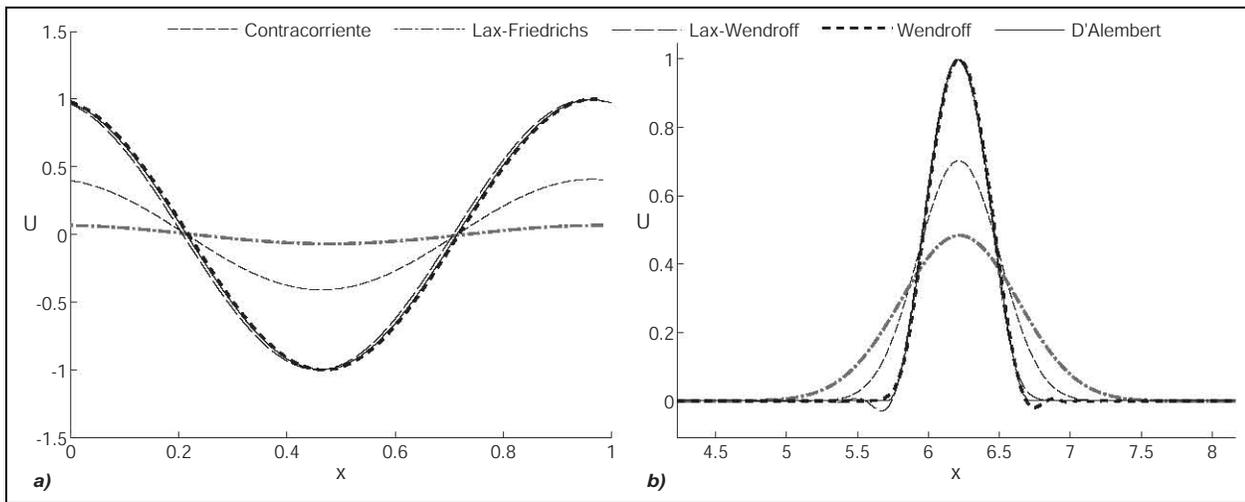


Figura 37: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 63$, $j = 721$ respectivas a condiciones iniciales: a) periódicas, y b) de soporte compacto.

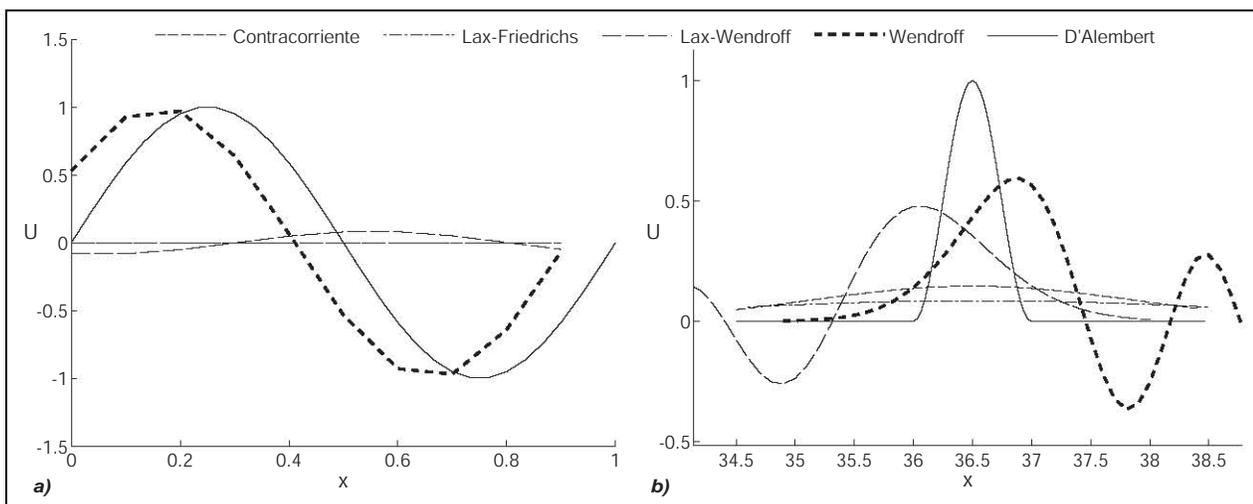


Figura 38: Gráficas de propagación para los diferentes esquemas numéricos, usando $C = 0.5$, $i_p = 10$, $j = 721$ respectivas a condiciones iniciales: a) periódicas, y b) de soporte compacto.

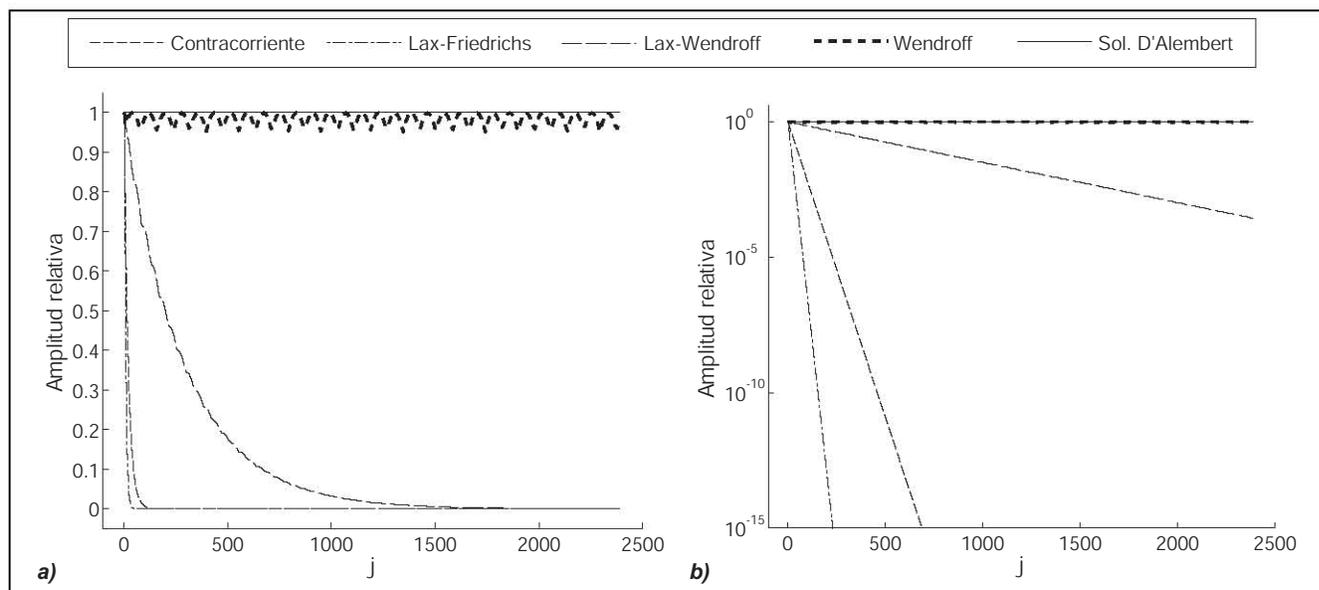


Figura 39: Amplitud relativa como función de j para condiciones iniciales periódicas donde $C = 0.5$, $i_p = 10$, y usando para A una escala a) lineal y b) logarítmica.

5.6.2. Amplitud relativa con condiciones iniciales periódicas.

Al graficar la amplitud relativa A contra j se observa difusión numérica en tres de los esquemas que hemos usado, que son precisamente los esquemas explícitos: contracorriente, Lax-Friedrichs y Lax-Wendroff. Para todas las combinaciones posibles de i_p y C que permiten apreciar difusión numérica se verifica que el esquema más difusivo es el de Lax-Friedrichs, en segundo lugar contracorriente y en tercer lugar Lax-Wendroff. En cuanto al esquema de Wendroff no es posible en ningún caso observar un decrecimiento de A , lo cual no es en sí argumento suficiente para decir que no exista difusión numérica; su total ausencia será justificada mediante otro tipo de argumentos. Cuando se usa un eje logarítmico para A , como en la figura 39b, puede apreciarse que en todos los casos su decrecimiento es una función exponencial con respecto a j ; es decir, se observan líneas rectas.

El hecho de que en ningún caso sea posible observar que decrezca la amplitud correspondiente al esquema de Wendroff, y de que para los otros esquemas decrezca exponencialmente es consistente con lo que se ha obtenido previamente al hacer la prueba de Von Neumann para cada uno de estos esquemas. En los primeros tres casos, contracorriente, Lax-Friedrichs, y Lax-Wendroff, se ha concluido que con cada paso temporal decrece la amplitud de los modos de Fourier correspondientes,

$$\left| \frac{\xi_m^{j+1}}{\xi_m^j} \right| < 1,$$

donde $\left| \xi_m^j \right|$ es la amplitud del m -ésimo modo correspondiente al nivel temporal t^j . Así mismo, en el caso del esquema de Wendroff este cociente siempre resultó ser igual a 1; es decir, no decrece la amplitud de sus modos de Fourier. En el presente caso de condiciones iniciales periódicas, como el perfil es senoidal, la propagación tiene solamente un modo de Fourier.

El parámetro i_p tiene una relación directa con A : en los esquemas explícitos y para los tres valores aplicables de C se observa que mientras más alto sea el valor de i_p , el decrecimiento de la amplitud es menor, y viceversa. Un ejemplo de las gráficas para Lax-Friedrichs con $C = 0.5$ se muestra en la figura 40.

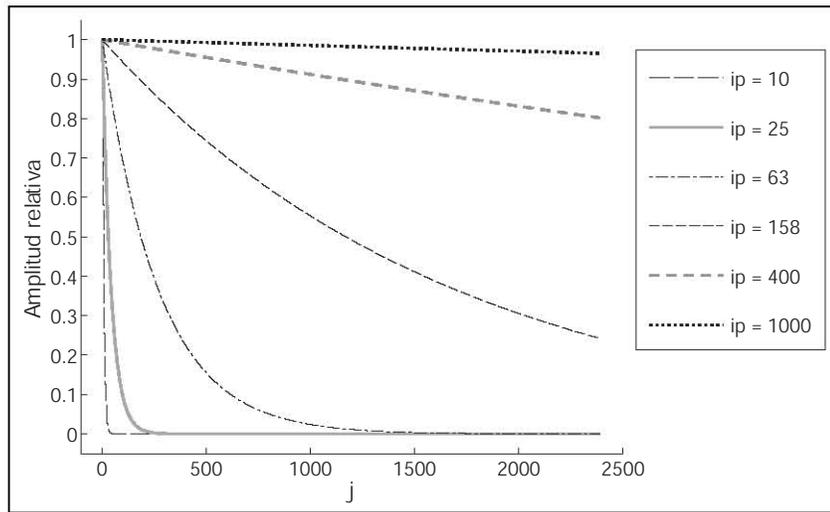


Figura 40: Amplitud relativa como función de j para Lax-Friedrichs con $C = 0.5$.

Con respecto a C la historia es parecida, pero no igual: independientemente del valor de i_p , para los esquemas contracorriente y Lax-Wendroff, $C = 0.5$ implica un decrecimiento de A más rápido que el que implica $C = 0.1$, y éste a su vez uno más rápido que el correspondiente a $C = 0.01$. Un ejemplo se muestra en la figura 41a. Sin embargo, para el esquema de Lax-Friedrichs se observa completamente lo opuesto; el caso $i_p = 158$ puede encontrarse en la figura 41b.

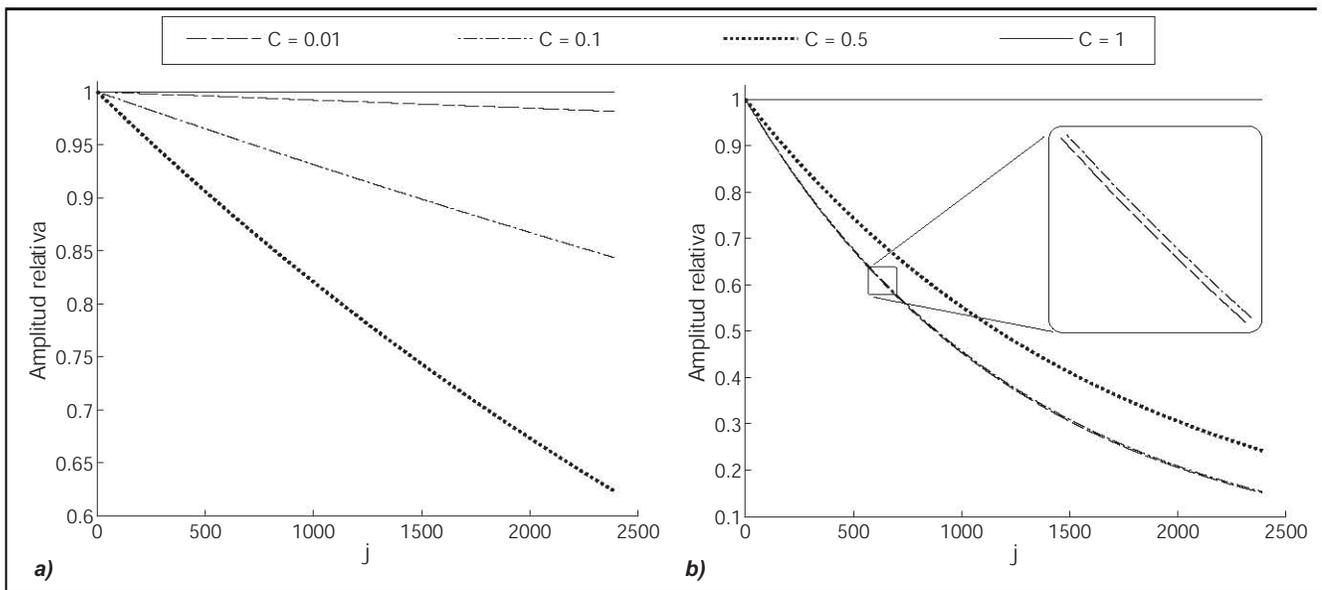


Figura 41: Amplitud relativa como función de j donde $i_p = 158$ para los esquemas a) contracorriente y b) Lax-Friedrichs.

5.6.3. Desplazamiento con condiciones iniciales periódicas.

Con respecto al desplazamiento D , observamos que los cuatro esquemas presentan dispersión numérica, sólo que es más fácil apreciarla para Lax-Wendroff y Wendroff que para los otros dos esquemas; sólo algunas combinaciones de i_p y C nos permiten apreciar dispersión para los esquemas contracorriente y de Lax-Friedrichs. Esto se debe

a que en estos esquemas las combinaciones que presentan mayor dispersión numérica también presentan mayor difusión numérica y por lo tanto rápidamente se pierde la resolución necesaria para medir D , como puede observarse por ejemplo en las figuras 32a y 38a. En las gráficas de la figura 43 las series correspondientes a Lax-Friedrichs están truncadas por esta razón. En todos los casos se observa aproximadamente una relación lineal entre j y D , lo cual indica que si bien las soluciones numéricas viajan a diferentes velocidades con diferentes parámetros, estas velocidades son constantes.

La magnitud del desplazamiento en todos los casos que se ha podido observar es mayor para valores pequeños de i_p y viceversa. Los casos que no se han podido observar son precisamente los valores más altos de i_p (porque serían necesarios muchos más pasos temporales para determinar que $D \neq 0$), en la mayoría de los caso a partir de $i_p = 63$. Un ejemplo se muestra en la figura 42.

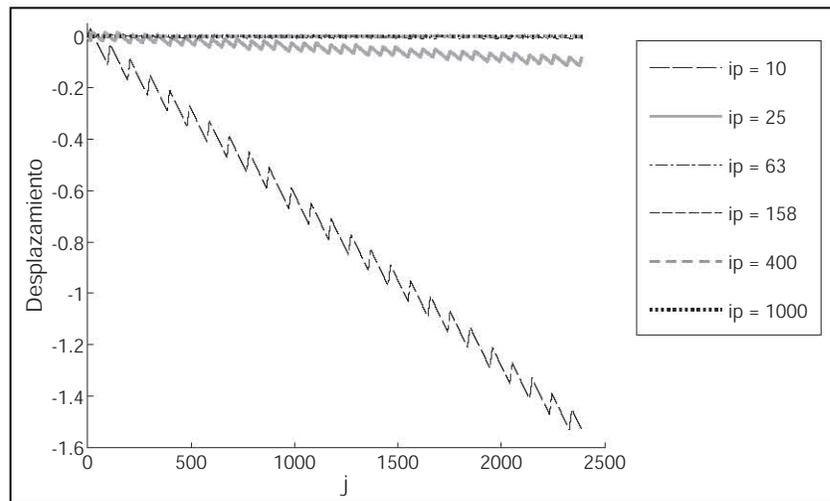


Figura 42: Desplazamiento como función de j para el esquema de Lax-Wendroff con $C = 0.1$.

Con respecto a C no es tan directa la relación, ya que los esquemas de Lax-Friedrichs y Lax-Wendroff son numéricamente menos dispersivos para $C = 0.1$ que para $C = 0.5$ y lo son aún menos para $C = 0.01$. El esquema contracorriente parece no tener un comportamiento monótono: por un lado es numéricamente más dispersivo para $C = 0.1$ que para $C = 0.01$ pero para $C = 0.5$ aparenta ser completamente no dispersivo. El caso del esquema de Wendroff muestra un comportamiento similar al de los dos primeros en el rango $C < 1$, pero por otra parte presenta mayor dispersión numérica para $C = 5$ que para $C = 1.1$. Estos datos indican que la relación entre C y D no es en todos los casos monótona y para poder apreciarla con detalle sería necesario efectuar cálculos utilizando una gama más amplia de valores de C .

Se observa en todos los casos tales que $C < 1$ que el esquema de Lax-Friedrichs es el que presenta una mayor dispersión numérica, en segundo lugar Lax-Wendroff, y tanto Wendroff como contracorriente son menos dispersivos que los dos anteriores —en algunos casos contracorriente es más dispersivo que Wendroff, y en otros casos se observa lo contrario, como se puede observar en la figura 43.

Otro aspecto que se puede apreciar en estas gráficas es que los desplazamientos se dan hacia ambos lados. Los esquemas contracorriente y Lax-Wendroff se retrasan respecto a la solución de D'Alembert, mientras que el esquema de Lax-Friedrichs se adelanta. El desplazamiento del esquema de Wendroff depende de la magnitud de C , para $C < 1$ se adelanta y para $C > 1$ se atrasa.

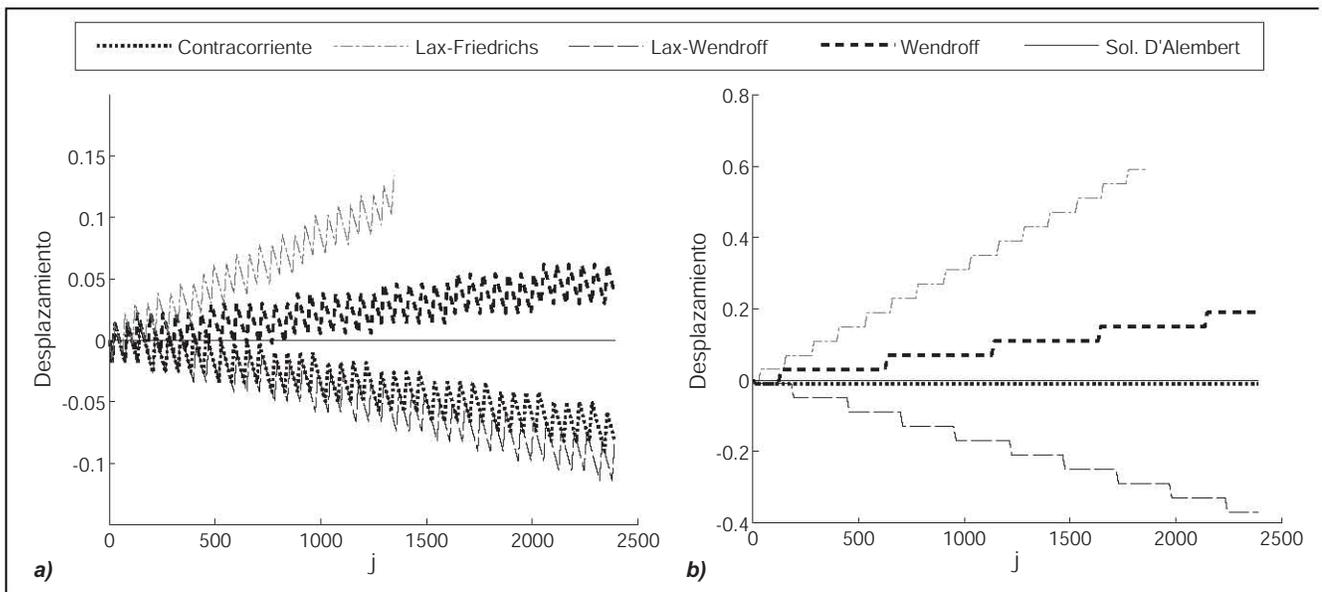


Figura 43: Desplazamiento como función de j para: a) $i_p = 25, C = 0.1$ y b) $i_p = 25, C = 0.5$.

5.6.4. Norma con condiciones iniciales periódicas.

La norma N es un parámetro que de alguna manera combina a los dos parámetros anteriores; si una propagación numérica presenta difusión numérica esperamos que su norma llegue eventualmente al valor 1 y se mantenga ahí, porque este sería el resultado obtenido cuando la amplitud relativa llega precisamente a cero. Por otro lado, si la propagación solamente presenta dispersión numérica esperamos que la norma oscile entre 0 y 2 porque al viajar a velocidades distintas la solución de D'Alembert y la solución numérica, alternarán entre estados donde se encuentren en fase y fuera de ella. El periodo de la oscilación nos dirá entonces que tan grande es la diferencia de velocidades; es decir, qué tan dispersivo es el esquema.

Siguiendo este razonamiento, lo que deberíamos observar en el caso de los esquemas explícitos es una combinación de ambas cosas: una oscilación cuya amplitud decrece y tiende hacia $N = 1$. Sin embargo, en los casos contracorriente y Lax-Friedrichs la amplitud decrece tan rápido que no es posible observar dicha oscilación. Por otra parte, el esquema implícito de Wendroff describe solamente oscilaciones, de modo que reafirma lo que se había mencionado: con los datos disponibles no se observa difusión numérica en este caso. Dichos resultados pueden observarse en la figura 44.

Por la presencia de las oscilaciones no podemos establecer directamente cuáles combinaciones de parámetros son mejores que otros en términos de la magnitud de sus normas. Aún así, tomando en cuenta las oscilaciones y la tendencia asintótica de N , en todos los casos que permiten observarlo (sobre todo para los valores bajos de i_p , donde se cuenta con suficientes puntos para ver oscilaciones completas y/o tendencias asintóticas) se verifican las observaciones al respecto de dispersión numérica y difusión numérica mencionadas en las secciones anteriores. Algunos ejemplos se muestran en la figura 45.

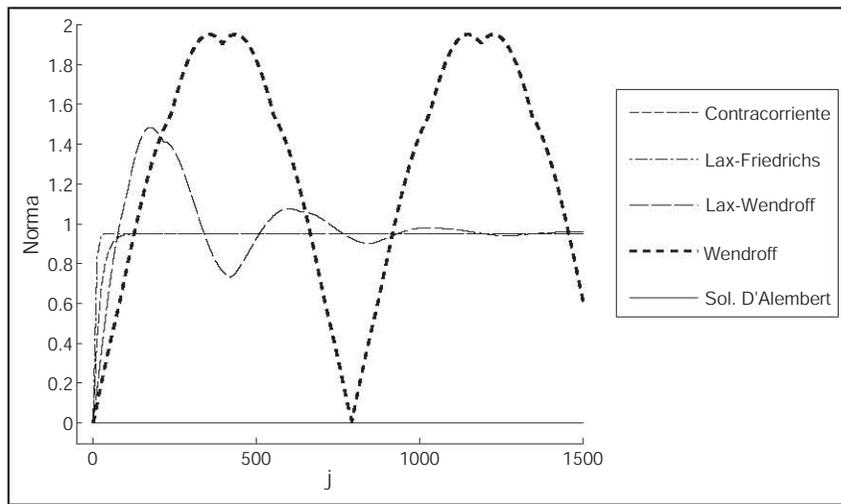


Figura 44: N como función de j para los diferentes esquemas numéricos con $i_p = 10$, $C = 0.5$.

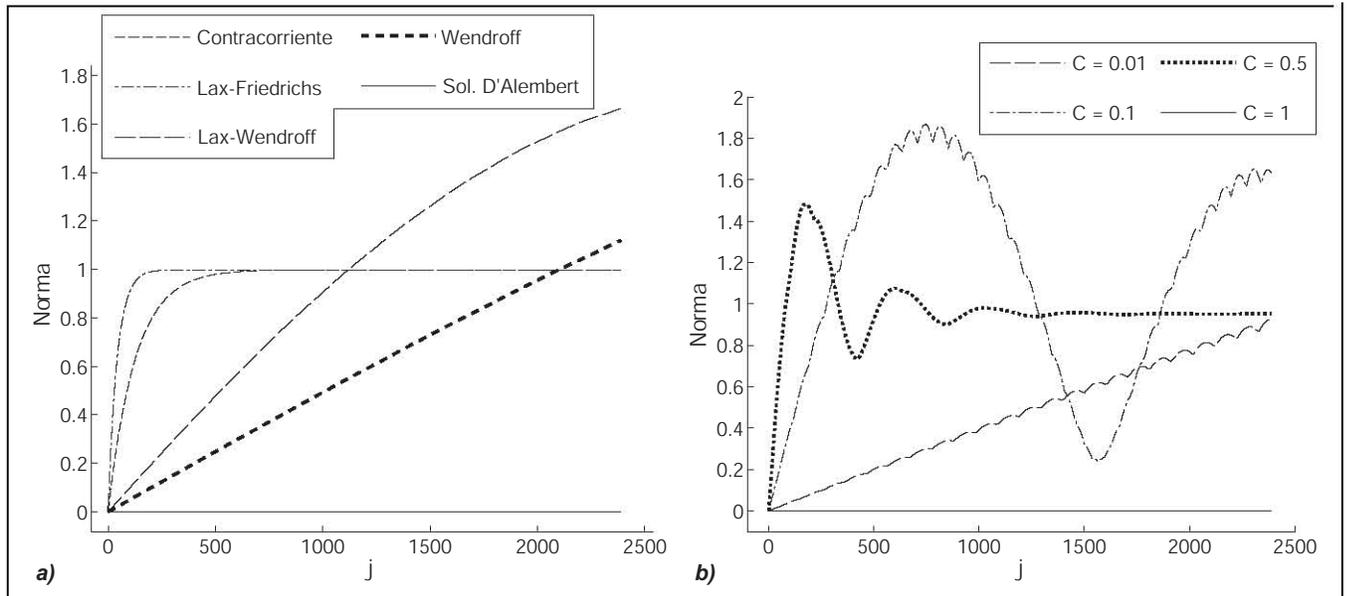


Figura 45: Norma como función de j para a) $C = 0.5$, $i_p = 25$ y los diferentes esquemas numéricos; y para b) $i_p = 10$ con el esquema de Lax-Wendroff.

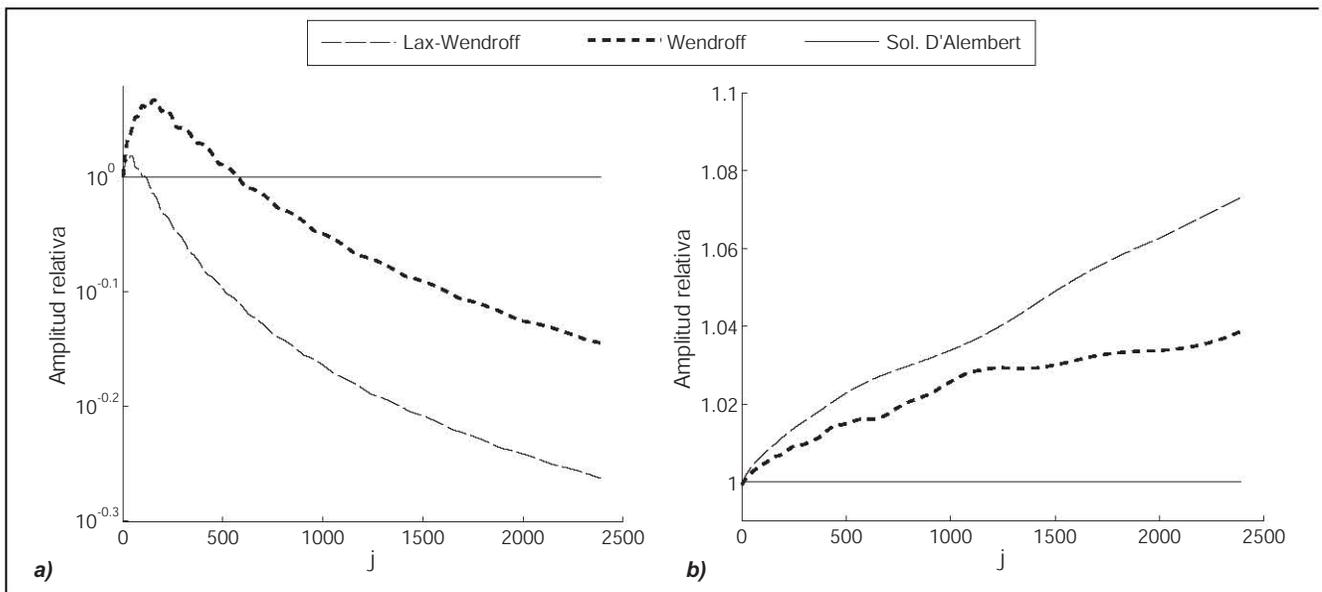


Figura 46: A como función de j para: a) $i_p = 10$, $C = 0.5$ y b) $i_p = 63$, $C = 0.5$. Comportamiento aparentemente anómalo para los esquemas de Lax-Wendroff y Wendroff.

5.6.5. Condiciones iniciales de soporte compacto.

Las observaciones a partir de los casos donde se ha tomado como condición inicial una función de soporte compacto son básicamente las mismas que para el caso periódico, es decir, se observan las mismas tendencias de los parámetros de evaluación A , D , N en relación a j , i_p y C . Sin embargo, hay algunas anomalías aparentes que a continuación se describen e interpretan a partir de las observaciones ya antes hechas sobre el caso de condiciones iniciales periódicas. La idea básica, tal como se ha mencionado en la sección 4.12.3 es que toda condición inicial de soporte compacto puede ser expresada mediante su transformada discreta de Fourier que es una sumatoria de funciones senoidales, para las cuales son válidas todas las observaciones que se han hecho en las secciones anteriores; es decir, la propagación de un pulso equivale a la suma de propagaciones de funciones senoidales.

La primera observación aparentemente anómala es el hecho de que para algunas combinaciones de C e i_p es evidente que decrece la amplitud del resultado obtenido por medio del esquema de Wendroff, como puede observarse en la figura 46a y, antes, en la figura 38b. Esto, en primera instancia, nos puede llevar a pensar en difusión numérica; sin embargo, esto entraría en contradicción con lo discutido al respecto de difusión numérica en las secciones anteriores, y por lo tanto concluimos que este efecto es debido a dispersión numérica. Las componentes senoidales (modos) del pulso, al ser propagadas mediante el esquema de Wendroff, presentan dispersión numérica: es decir, deben propagarse a velocidades diferentes, lo que ocasiona un desfase entre ellas, y esto necesariamente traerá consigo alguna deformación de la señal, que bien puede representar un aumento o disminución de la amplitud. El aumento de la amplitud es otro comportamiento aparentemente anómalo, que se muestra de modo aún más claro en la figura 46b para el esquema de Wendroff e incluso para el esquema de Lax-Wendroff. En los esquemas de Lax-Friedrichs y contracorriente, que también son numéricamente dispersivos, no podemos observar un aumento de su amplitud porque para estos esquemas la difusión numérica es más pronunciada.

Existe también una manifestación anómala para los esquemas de contracorriente y Lax-Friedrichs: en la figura 47 se ha usado un eje logarítmico para A pero, a diferencia de lo que podía observarse en el caso periódico, no

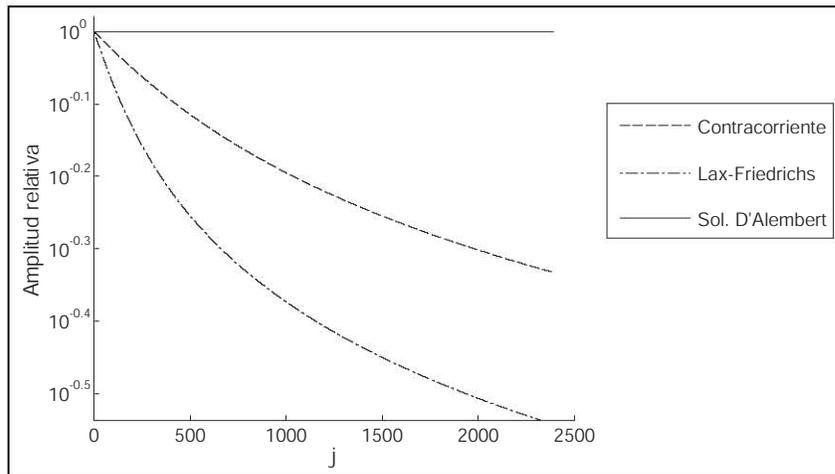


Figura 47: A como función de j para $i_p = 63$, $C = 0.5$. Comportamiento aparentemente anómalo de los esquemas contracorriente y Lax-Friedrichs.

son líneas rectas las que describen el cambio en la magnitud de A ; es decir, esta curva no es exponencial. Este comportamiento puede explicarse de nuevo pensando en la condición inicial de soporte compacto como una suma de funciones senoidales, cuya amplitud decrece exponencialmente, pero para unos modos más rápido que para otros, y por lo tanto el decrecimiento la amplitud de la suma no tiene por qué ser exponencial.

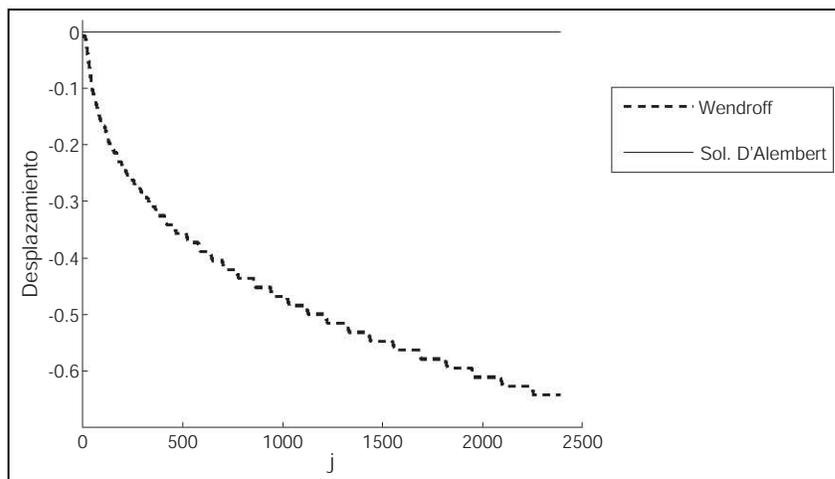


Figura 48: Desplazamiento como función de j para $C = 5$, $i_p = 65$. Comportamiento aparentemente anómalo del esquema de Wendroff.

Con respecto al desplazamiento D , observamos cómo, en la figura 48, para el esquema de Wendroff parece no haber una relación lineal entre D y j . Esta anomalía aparente también se explica por medio de la dispersión numérica asociada con este esquema. Cada modo efectivamente se propaga a una velocidad constante, pero diferente, deformando así el perfil de la señal, por lo que su máximo (que es el punto usado para medir el desplazamiento D) no necesariamente debe desplazarse a una velocidad constante. Para los otros esquemas se pueden hacer observaciones similares, donde D no guarda una relación lineal con j , pero donde gráficamente son menos evidentes estos efectos.

5.6.6. Resumen de observaciones.

En primer lugar observamos que los esquemas de segundo orden, Lax-Wendroff y Wendroff representan en todas las circunstancias notablemente una mejor aproximación a la solución de D'Alembert que los esquemas de primer orden, contracorriente y Lax-Friedrichs.

En términos generales la mejor aproximación la brinda el esquema de Wendroff, en segundo lugar el de Lax-Wendroff, luego contracorriente y por último el de Lax-Friedrichs. Las excepciones dependen del parámetro que se use para comparar la proximidad, por ejemplo en términos de desplazamiento D la figura 43 nos muestra que el esquema contracorriente es mejor que el esquema Lax-Wendroff, y en un caso, mejor que el esquema de Wendroff. Se ha verificado también, tanto teórica como numéricamente, que el esquema de Wendroff no presenta difusión numérica, y esta es una ventaja sobre los otros tres esquemas.

En cuanto a los parámetros i_p y C , entre ellos existe una diferencia esencial. Todos los esquemas obtienen mejores resultados siempre que el valor de i_p es mayor. Sin embargo, en cuanto al valor del parámetro C , en términos de amplitud relativa A , en la mayoría de los casos se obtiene un mejor resultado cuando C es más pequeño, pero en el caso de Lax-Friedrichs es precisamente lo contrario. Por otro lado, en términos del desplazamiento D , el valor de C siempre brinda mejores resultados cuando es igual a 0.01 (comparado con los otros dos valores disponibles) pero en el caso del esquema contracorriente (y también en caso de Wendroff si se consideran los casos $C > 1$) no exhibe una relación monótona con respecto a D .

Adicionalmente concluimos que es más conveniente hacer la caracterización de los esquemas numéricos a partir de observaciones hechas sobre propagaciones del tipo senoidal, las cuales nos permiten posteriormente interpretar la propagación numérica de perfiles más complicados.

Como una ilustración de los resultados expuestos se muestra una última serie de gráficas donde se han usado los valores óptimos de C e i_p para cada uno de los esquemas, descartando el caso $C = 1$ para el cual se ha mostrado en las secciones 5.1, 5.2, 5.3 y 5.4 que los resultados numéricos corresponden precisamente a la solución exacta de la ecuación de advección, es decir, se han usado los valores $i_p = 1000$ y $C = 0.01$ para los esquemas contracorriente, de Lax-Wendroff, y de Wendroff; y los valores $i_p = 1000$ y $C = 0.5$ para el esquema de Lax-Friedrichs.

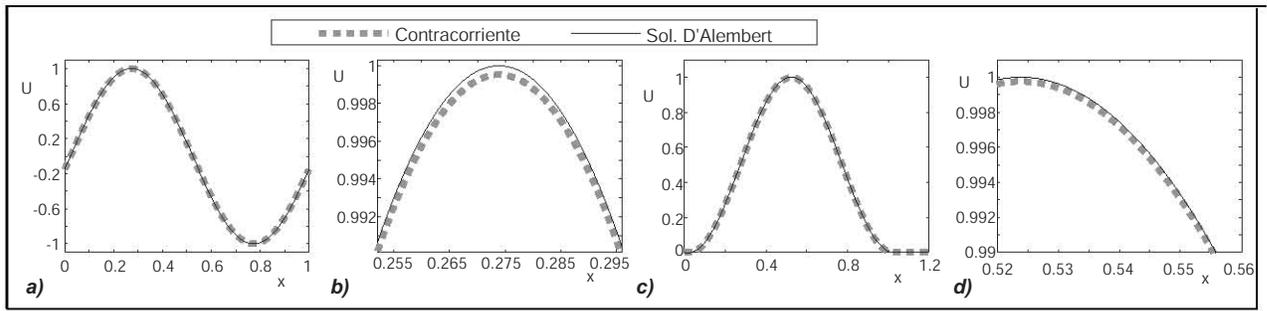


Figura 49: Resultados numéricos para el esquema contracorriente con $j = 2389$, $i_p = 1000$, $C = 0.01$. El caso **a)** corresponde a condiciones iniciales periódicas y el caso **c)** a condiciones iniciales de soporte compacto, los casos **b)** y **d)** son detalles de los casos anteriores, respectivamente.

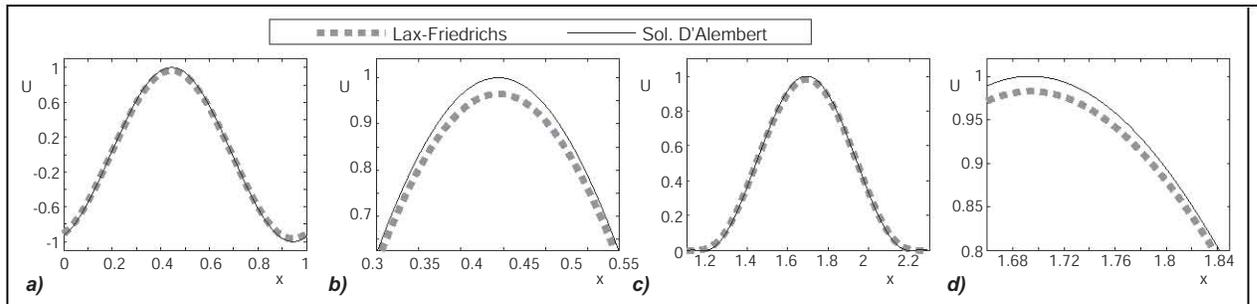


Figura 50: Resultados numéricos para el esquema de Lax-Friedrichs con $j = 2389$, $i_p = 1000$, $C = 0.5$. El caso **a)** corresponde a condiciones iniciales periódicas y el caso **c)** a condiciones iniciales de soporte compacto, los casos **b)** y **d)** son detalles de los casos anteriores, respectivamente.

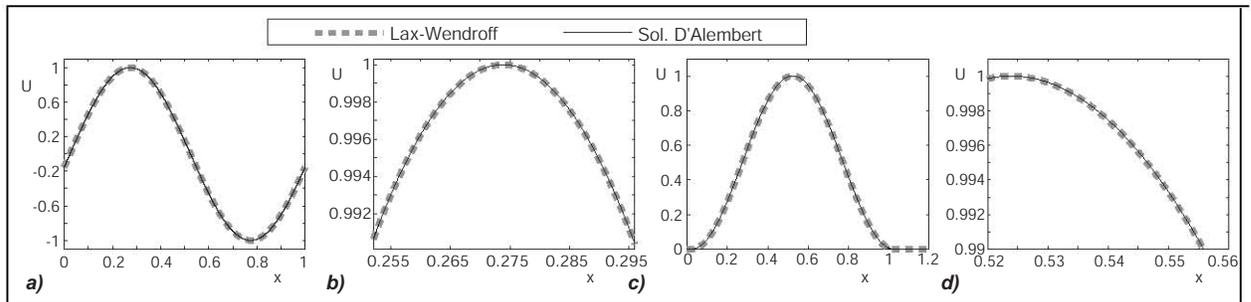


Figura 51: Resultados numéricos para el esquema de Lax-Wendroff con $j = 2389$, $i_p = 1000$, $C = 0.01$. El caso **a)** corresponde a condiciones iniciales periódicas y el caso **c)** a condiciones iniciales de soporte compacto, los casos **b)** y **d)** son detalles de los casos anteriores, respectivamente.

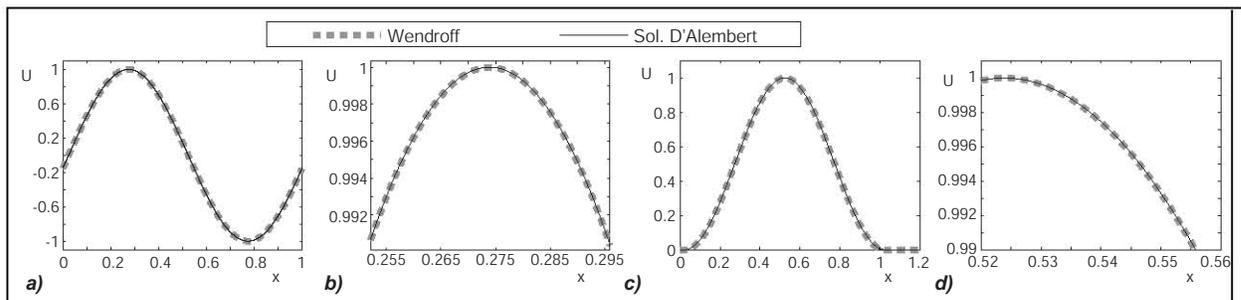


Figura 52: Resultados numéricos para el esquema de Wendroff con $j = 2389$, $i_p = 1000$, $C = 0.01$. El caso **a)** corresponde a condiciones iniciales periódicas y el caso **c)** a condiciones iniciales de soporte compacto, los casos **b)** y **d)** son detalles de los casos anteriores, respectivamente.

6. Ecuaciones modificadas para esquemas numéricos de la ecuación de advección.

Una de las facilidades que presenta la ecuación de advección es que conocemos su solución explícitamente y podemos compararla en todo momento con las soluciones numéricas. Como hemos visto a lo largo de la sección 5.6, cuando calculamos por medio de alguno de los esquemas una solución numérica U , obtenemos una aproximación a u que en algunos casos presenta algunas cualidades que no pueden observarse en la misma u , ya que el perfil de la solución numérica U se deforma mientras viaja. Desde luego atribuimos este fenómeno al hecho de que U está construida sobre un dominio discreto, es decir, a su carácter numérico, pero ¿podemos predecir cual será la deformación que corresponde a cada uno de los métodos?. En cierta medida sí podemos hacerlo, y para ello encontraremos una segunda función \tilde{u} que se asemeja mucho a la solución numérica U .

Si observamos el modo como hemos calculado la magnitud del error de truncamiento, en las ecuaciones (80), (89), (96) y (103)-(106) se han sustraído algunos términos gracias a que u satisface la ecuación de advección. Pero, ¿qué pasaría si u fuera solución de una ecuación diferencial diferente?. En este caso sería posible que pudiéramos sustraer menos términos, o más. Lo interesante es precisamente esto último: para cada uno de los esquemas es posible encontrar varias ecuaciones diferenciales cuyas soluciones, al ser sustituidas en las ecuaciones (80), (89), (96) y (103)-(106) eliminan un mayor número de términos. Esto significa dos cosas: por un lado que el esquema numérico tendría también que ser consistente con esas ecuaciones, y en segundo lugar que el error de truncamiento del esquema numérico con respecto a estas nuevas ecuaciones sería menor que el respectivo a la ecuación de advección. En otras palabras, cada esquema numérico sería de orden mayor con respecto a estas ecuaciones que con respecto a la ecuación de advección. Para cada esquema escogeremos una de estas nuevas ecuaciones, cuya solución \tilde{u} podamos conocer explícitamente, y la llamaremos *ecuación modificada*, en inglés *modified equation* [33] o en algunas ocasiones también llamada *model equation* [35].

Para los esquemas numéricos de la ecuación de advección que se han expuesto siempre será posible escoger la ecuación modificada en alguna de las siguientes formas:

$$\tilde{u}_t + c\tilde{u}_x = \alpha\tilde{u}_{xx}, \quad (114)$$

$$\tilde{u}_t + c\tilde{u}_x = \mu\tilde{u}_{xxx}. \quad (115)$$

La primera de estas ecuaciones se llama *ecuación advección-difusión* y la segunda es una ecuación dispersiva, a la cual en adelante llamaremos *ecuación de advección con dispersión*. En las secciones 6.1 y 6.2, respectivamente, se muestran desarrollos para obtener sus soluciones exactas.

Para cada una de las dos alternativas de ecuaciones modificadas que se han presentado hay un parámetro que determina la magnitud de la deformación de la onda. Como se mostrará en las secciones siguientes, dichos parámetros α y μ son siempre funciones de Δx , Δt , y su forma explícita dependerá del esquema numérico que se esté usando. Esto significa que podemos caracterizar los esquemas numéricos mediante el nivel y el tipo de deformación que implican ya sea α o μ , y además el nivel de deformación del perfil de U dependerá del nivel de refinamiento de la malla.

6.1. Solución de la ecuación de advección-difusión

Seguimos el procedimiento expuesto por Le Veque [35]. La ecuación de advección-difusión (114) es

$$\tilde{u}_t + c\tilde{u}_x = \alpha\tilde{u}_{xx}.$$

Ahora definimos una función v que tenga la siguiente propiedad

$$v(\eta, t) = \tilde{u}(\eta + ct, t),$$

donde $\eta = x - ct$, por lo que

$$\begin{aligned}\tilde{u}(x, t) &= v(\eta, t), \\ \tilde{u}_x &= v_\eta, \\ \tilde{u}_{xx} &= v_{\eta\eta}, \\ \tilde{u}_t &= (-c)v_\eta + v_t.\end{aligned}$$

Entonces v debe cumplir con la ecuación

$$(-c)v_\eta + v_t + cv_\eta = \alpha v_{\eta\eta}.$$

Es decir,

$$v_t = \alpha v_{\eta\eta},$$

que es precisamente la ecuación de calor. Una solución de esta ecuación es

$$v(\eta, t) = \exp(-\alpha\omega^2 t) \cdot (A \sin(\omega\eta) + B \cos(\omega\eta)),$$

donde ω es cualquier número real. Como esta ecuación es homogénea la solución general se puede expresar para intervalos infinitos como una integral sobre ω donde los coeficientes A y B se convierten en funciones de ω , o bien, para intervalos finitos o condiciones iniciales periódicas como una serie,

$$v(\eta, t) = \sum_{k=0}^{\infty} \exp\left(-\alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 t\right) \left(A_k \sin\left(\frac{2\pi k}{\lambda_M}\eta\right) + B_k \cos\left(\frac{2\pi k}{\lambda_M}\eta\right)\right),$$

donde λ_M es la longitud de onda o la longitud del intervalo y k es un entero.

Cuando $t = 0$ se tiene $v = \tilde{u}$, y la serie es precisamente una serie de Fourier por lo que los coeficientes A_k y B_k quedan unívocamente determinados al satisfacer la condición inicial.

Entonces la solución buscada de la ecuación de advección-difusión, para condiciones iniciales periódicas o intervalos finitos, es:

$$\tilde{u}(x, t) = \sum_{k=0}^{\infty} \exp\left(-\alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 t\right) \left(A_k \sin\left(\frac{2\pi k}{\lambda_M}(x - ct)\right) + B_k \cos\left(\frac{2\pi k}{\lambda_M}(x - ct)\right)\right).$$

Ahora tomemos uno solo de los términos de esta suma

$$\exp\left(-\alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 t\right)\left(A_k \sin\left(\frac{2\pi k}{\lambda_M}(x-ct)\right) + B_k \cos\left(\frac{2\pi k}{\lambda_M}(x-ct)\right)\right),$$

y aprovechamos las propiedades de las funciones trigonométricas para expresarlo del siguiente modo:

$$\exp\left(-\alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 t\right) \cdot D_k \sin\left(\frac{2\pi k}{\lambda_M}(x-ct) + \phi_k\right).$$

Esta es una función senoidal que se desplaza hacia la derecha con una velocidad c y cuya amplitud decrece exponencialmente. Notamos además que el factor de decrecimiento

$$\tau_k = \alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 \quad (116)$$

está relacionado a través de k con la longitud de onda, en el sentido de que si la longitud de onda es menor el decrecimiento es más rápido. Eso quiere decir que si la solución correspondiente a nuestra condición inicial consiste de varios modos de propagación, entonces observaremos una deformación de la señal mientras viaja y no solamente una disminución de la amplitud. Concretamente el perfil tenderá a hacerse menos puntiagudo y más redondo.

Entonces, es posible expresar la solución de la ecuación de advección-difusión como:

$$\tilde{u}(x, t) = \sum_{k=0}^{\infty} \exp\left(-\alpha\left(\frac{2\pi k}{\lambda_M}\right)^2 t\right) \cdot D_k \sin\left(\frac{2\pi k}{\lambda_M}(x-ct) + \phi_k\right). \quad (117)$$

Este resultado es suficiente para nosotros en el sentido de que nos permitirá implementar las ecuaciones modificadas y compararlas con los resultados numéricos.

6.2. Solución de la ecuación de advección con dispersión.

La ecuación modificada (115) correspondiente a los esquemas Wendroff y Lax-Wendroff es una ecuación dispersiva del siguiente tipo:

$$\tilde{u}_t + c\tilde{u}_x = \mu\tilde{u}_{xxx}, \quad (118)$$

a la cual llamaremos *ecuación de advección con dispersión*. Como esta es una ecuación homogénea entonces buscaremos su solución en la forma de una exponencial compleja del tipo

$$A \exp(i2\pi(\gamma x + \nu t)),$$

y al sustituir en la ecuación (118) y eliminar los factores que aparecen en todos los términos se obtiene

$$\nu + c\gamma = -\mu(2\pi)^2 \gamma^3,$$

de modo que

$$\nu = -(c\gamma + \mu(2\pi)^2 \gamma^3).$$

Entonces la función

$$A \exp\left(i2\pi\left(\gamma x - \left(c\gamma + \mu(2\pi)^2 \gamma^2\right)t\right)\right)$$

es una solución de la ecuación (118). Para obtener la solución general en términos de la condición inicial

$$\tilde{u}(x, 0) = g(x) ,$$

observamos que si consideramos A como una función de γ , cuando $t = 0$, podemos integrar sobre γ de $-\infty$ a ∞ y resulta que $A(\gamma)$ es precisamente la transformada de Fourier de $g(x)$, y entonces

$$\tilde{u} = \int_{-\infty}^{\infty} A(\gamma) \exp\left(i2\pi\left(\gamma x - \left(c\gamma + \mu(2\pi)^2 \gamma^2\right)t\right)\right) d\gamma .$$

Cuando $g(x)$ es una función periódica o definida sobre un intervalo finito podemos expresar la solución \tilde{u} como una serie,

$$\tilde{u} = \sum_{k=0}^{\infty} \left(A_k \cos\left(\frac{2\pi}{\lambda_M} kx - \left(\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3\right)t\right) + B_k \sin\left(\frac{2\pi}{\lambda_M} kx - \left(\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3\right)t\right) \right) ,$$

donde k es un entero, λ_M es la longitud de onda o la longitud del intervalo y los coeficientes A_k, B_k son precisamente los coeficientes de la serie de Fourier de g que se obtiene cuando $t = 0$.

Nuevamente para establecer las propiedades de esta solución tomamos uno de los términos de la serie y aprovechamos las propiedades de las funciones trigonométricas para expresarla como una sola función senoidal,

$$D_k \sin\left(\frac{2\pi}{\lambda_M} kx - \left(\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3\right)t + \phi_k\right) .$$

Lo importante en este caso es que esta es una función senoidal que se desplaza sin deformarse hacia la derecha con una velocidad igual al inverso aditivo del cociente de los factores que acompañan a x y a t

$$\begin{aligned} \tilde{c}_k &= \frac{\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3}{\frac{2\pi}{\lambda_M} k} \\ &= c + \mu \frac{(2\pi)^2}{\lambda_M^2} k^2 , \end{aligned} \quad (119)$$

donde dicha velocidad está relacionada con la longitud de onda por medio de k ; es decir, cada uno de los modos de propagación se desplazará a una velocidad diferente, y de este modo la onda se deforma mientras viaja.

Podemos expresar la solución entonces como:

$$\tilde{u} = \sum_{k=0}^{\infty} D_k \sin\left(\frac{2\pi}{\lambda_M} kx - \left(\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3\right)t + \phi_k\right) ; \quad (120)$$

del mismo modo que en el caso anterior, este resultado nos permitirá implementar las ecuaciones modificadas y compararlas con los resultados numéricos.

6.3. Expresiones derivadas de las ecuaciones diferenciales.

El hecho de que \tilde{u} satisfaga alguna de las ecuaciones (114) o (115) será aprovechado en secciones posteriores sustituyendo, por ejemplo en el caso del esquema contracorriente, las apariciones de \tilde{u}_t por los términos $\alpha\tilde{u}_{xx} - c\tilde{u}_x$. Sin embargo, bajo el supuesto de que \tilde{u} es suficientemente suave como para que sea posible intercambiar el orden de sus segundas y terceras derivadas, es posible derivar todos los términos de dichas ecuaciones diferenciales y obtener de este modo otras identidades que nos permitirán hacer sustituciones semejantes. Algunas de dichas identidades se encuentran listadas a continuación:

- Si \tilde{u} satisface la ecuación de advección-difusión (114)

$$\tilde{u}_t + c\tilde{u}_x = \alpha\tilde{u}_{xx} ,$$

entonces

$$\begin{aligned} \tilde{u}_t &= -c\tilde{u}_x + \alpha\tilde{u}_{xx} , \\ \tilde{u}_{tx} &= -c\tilde{u}_{xx} + \alpha\tilde{u}_{xxx} , \\ \tilde{u}_{tt} &= c^2\tilde{u}_{xx} - c\alpha\tilde{u}_{xxx} + \alpha\tilde{u}_{xxt} . \end{aligned} \tag{121}$$

- Si \tilde{u} satisface la ecuación de advección con dispersión (115)

$$\tilde{u}_t + c\tilde{u}_x = \mu\tilde{u}_{xxx} ,$$

entonces

$$\begin{aligned} \tilde{u}_t &= -c\tilde{u}_x + \mu\tilde{u}_{xxx} , \\ \tilde{u}_{tx} &= -c\tilde{u}_{xx} + \mu\tilde{u}_{xxx} , \\ \tilde{u}_{tt} &= c^2\tilde{u}_{xx} + \mu(-c\tilde{u}_{xxx} + \tilde{u}_{xxt}) , \\ \tilde{u}_{txx} &= -c\tilde{u}_{xxx} + \mu\tilde{u}_{xxxx} , \\ \tilde{u}_{ttx} &= c^2\tilde{u}_{xxx} + \mu(-c\tilde{u}_{xxxx} + \tilde{u}_{xxxxt}) , \\ \tilde{u}_{ttt} &= -c^3\tilde{u}_{xxx} + \mu(c^2\tilde{u}_{xxxx} - c\tilde{u}_{xxxxt} + \tilde{u}_{xxxxt}) . \end{aligned} \tag{122}$$

6.4. Valor de α para el esquema contracorriente

Para obtener la ecuación modificada de este esquema contracorriente usamos la expresión (80)

$$L = u_t + cu_x + \Delta t \frac{1}{2} u_{tt} - \Delta x \frac{c}{2} u_{xx} + O((\Delta t)^2) + O((\Delta x)^2) ,$$

donde ya hemos observado que si u satisface la ecuación de advección podemos eliminar los términos de orden cero y el error de truncamiento tiene únicamente términos de primer orden. Si sustituimos u por \tilde{u} solución de la ecuación (114),

$$\tilde{u}_t + c\tilde{u}_x = \alpha\tilde{u}_{xx} ,$$

entonces podremos eliminar más términos y el error de truncamiento tendrá únicamente términos de segundo orden y superiores. Para hacerlo notamos que la ecuación (114) implica las relaciones (121) expuestas en la sección 6.3.

Hacemos efectivamente la sustitución y obtenemos

$$\begin{aligned}
 L &= \tilde{u}_t + c\tilde{u}_x + \Delta t \frac{1}{2} \tilde{u}_{tt} - \Delta x \frac{c}{2} \tilde{u}_{xx} + O((\Delta t)^2) + O((\Delta x)^2) \\
 &= -c\tilde{u}_x + \alpha \tilde{u}_{xx} + c\tilde{u}_{xx} + \Delta t \frac{1}{2} (c^2 \tilde{u}_{xx} - c\alpha \tilde{u}_{xxx} + \alpha \tilde{u}_{xxx}) - \Delta x \frac{c}{2} \tilde{u}_{xx} + O((\Delta t)^2) + O((\Delta x)^2) \\
 &= \tilde{u}_{xx} \left(\alpha + \Delta t \frac{c^2}{2} - \Delta x \frac{c}{2} \right) - \Delta t \frac{c\alpha}{2} \tilde{u}_{xxx} + \Delta t \frac{\alpha}{2} \tilde{u}_{xxx} + O((\Delta t)^2) + O((\Delta x)^2) .
 \end{aligned}$$

Como puede observarse en esta expresión, para que el esquema numérico contracorriente sea de segundo orden con respecto a su ecuación modificada, según la definición expuesta en la sección 4.10, α debe cumplir

$$\alpha = \Delta x \frac{c}{2} - \Delta t \frac{1}{2} c^2 , \quad (123)$$

ya que de esta manera se anula el primer término, mientras que, bajo el supuesto de que al refinar la malla el cociente $\frac{\Delta t}{\Delta x}$ es constante, los términos posteriores se vuelven al menos de segundo orden con respecto a Δt . De modo equivalente podemos utilizar las ecuaciones (57) y (71) (o bien (73)):

$$\begin{aligned}
 C &= \frac{c \Delta t}{\Delta x} , \\
 i_p &= \frac{1}{\Delta x} ,
 \end{aligned}$$

para expresar α en términos de i_p y C , y se obtiene

$$\alpha = \frac{c}{2i_p} (1 - C) .$$

Entonces la ecuación modificada para este esquema numérico puede expresarse, por ejemplo, como:

$$\tilde{u}_t + c\tilde{u}_x = \left(\Delta x \frac{c}{2} - \Delta t \frac{1}{2} c^2 \right) \tilde{u}_{xx} ,$$

que es precisamente la ecuación de advección-difusión, cuya solución puede encontrarse en la sección 6.1.

6.5. Valor de α para el esquema de Lax-Friedrichs.

Al calcular el error de truncamiento del esquema de Lax-Friedrichs hemos obtenido la ecuación (89)

$$L = u_t + cu_x + \Delta t \frac{1}{2} u_{tt} - \frac{(\Delta x)^2}{2(\Delta t)} u_{xx} + O((\Delta t)^2) + \frac{1}{\Delta t} O((\Delta x)^3) + O((\Delta x)^2) ,$$

y ahora, para obtener la ecuación modificada se sustituye u por aquella \tilde{u} que satisface la ecuación (114), y por consiguiente las relaciones (121). Al hacer esta sustitución se obtiene:

$$\begin{aligned} L &= \tilde{u}_t + c\tilde{u}_x + \Delta t \frac{1}{2} \tilde{u}_{tt} - \frac{(\Delta x)^2}{2(\Delta t)} \tilde{u}_{xx} + O((\Delta t)^2) + \frac{1}{\Delta t} O((\Delta x)^3) + O((\Delta x)^2) \\ &= -c\tilde{u}_x + \alpha \tilde{u}_{xx} + c\tilde{u}_x + \Delta t \frac{1}{2} (c^2 \tilde{u}_{xx} - c\alpha \tilde{u}_{xxx} + \alpha \tilde{u}_{xxt}) - \frac{(\Delta x)^2}{2(\Delta t)} \tilde{u}_{xx} + O((\Delta t)^2) + O((\Delta x)^2) + \frac{1}{\Delta t} O((\Delta x)^3) \\ &= \tilde{u}_{xx} \left(\alpha + \frac{c^2}{2} \Delta t - \frac{(\Delta x)^2}{2(\Delta t)} \right) - \Delta t \frac{c\alpha}{2} \tilde{u}_{xxx} + \Delta t \frac{\alpha}{2} \tilde{u}_{xxt} + O((\Delta t)^2) + O((\Delta x)^2) \frac{1}{\Delta t} O((\Delta x)^3), \end{aligned}$$

de donde se observa que, para que este esquema numérico sea de segundo orden con respecto a su respectiva ecuación modificada (bajo la definición que se ha expuesto en la sección 4.10), α debe satisfacer

$$\alpha = \frac{(\Delta x)^2}{2(\Delta t)} - \frac{1}{2} (\Delta t) c^2, \quad (124)$$

y por lo tanto dicha ecuación modificada es la siguiente ecuación de advección-difusión:

$$\tilde{u}_t + c\tilde{u}_x = \left(\frac{(\Delta x)^2}{2(\Delta t)} - \frac{1}{2} (\Delta t) c^2 \right) \tilde{u}_{xx},$$

cuya solución puede encontrarse en la sección 6.1.

Podemos además utilizar las expresiones (57) y (71) (o bien (73)):

$$\begin{aligned} C &= \frac{c \Delta t}{\Delta x}, \\ i_p &= \frac{1}{\Delta x}, \end{aligned}$$

para expresar α en términos de C e i_p , al sustituir dichas ecuaciones en la expresión (124) se obtiene

$$\alpha = \frac{c}{2i_p} \left(\frac{1}{C} - C \right).$$

6.6. Valor de μ para el esquema Lax-Wendroff.

Para obtener la ecuación modificada del esquema de Lax-Wendroff partimos de la expresión (96)

$$L = u_t + cu_x + \Delta t \frac{1}{2} u_{tt} - \frac{c^2 (\Delta t)}{2} u_{xx} + (\Delta t)^2 \frac{1}{6} u_{ttt} + c (\Delta x)^2 \frac{1}{6} u_{xxx} + O((\Delta t)^3) + O((\Delta x)^3) + \Delta t O((\Delta x)^2),$$

y sustituimos u por \tilde{u} bajo el supuesto de que \tilde{u} satisface la ecuación (115),

$$\tilde{u}_t + c\tilde{u}_x = \mu \tilde{u}_{xxx},$$

y por lo tanto se cumplen las relaciones (122). Entonces, al efectuar dicha sustitución se tiene

$$\begin{aligned}
L &= \tilde{u}_t + c\tilde{u}_x + \Delta t \frac{1}{2} \tilde{u}_{tt} - \Delta t \frac{c^2}{2} \tilde{u}_{xx} + (\Delta t)^2 \frac{1}{6} \tilde{u}_{ttt} + (\Delta x)^2 \frac{c}{6} \tilde{u}_{xxx} \\
&\quad + O((\Delta t)^3) + O((\Delta x)^3) + \Delta t O((\Delta x)^2) \\
&= -c\tilde{u}_x + \mu \tilde{u}_{xxx} + c\tilde{u}_x + \Delta t \frac{1}{2} (c^2 \tilde{u}_{xx} + \mu (-c\tilde{u}_{xxx} + \tilde{u}_{xxt})) - \Delta t \frac{c^2}{2} \tilde{u}_{xx} \\
&\quad + (\Delta t)^2 \frac{1}{6} (-c^3 \tilde{u}_{xxx} + \mu (c^2 \tilde{u}_{xxxx} - c\tilde{u}_{xxx} + \tilde{u}_{xxt})) + (\Delta x)^2 \frac{c}{6} \tilde{u}_{xxx} \\
&\quad + O((\Delta t)^3) + O((\Delta x)^3) + \Delta t O((\Delta x)^2) \\
&= \tilde{u}_{xxx} \left(\mu - (\Delta t)^2 \frac{c^3}{6} + (\Delta x)^2 \frac{c}{6} \right) \\
&\quad + \Delta t \frac{\mu}{2} (-c\tilde{u}_{xxx} \tilde{u}_{xxt}) + (\Delta t)^2 \frac{\mu}{6} (c^2 \tilde{u}_{xxxx} - c\tilde{u}_{xxx} + \tilde{u}_{xxt}) \\
&\quad + O((\Delta t)^3) + O((\Delta x)^3) + \Delta t O((\Delta x)^2) ,
\end{aligned}$$

de modo que para que este esquema numérico sea de tercer orden con respecto a una ecuación modificada del tipo (115), siguiendo la definición expuesta en la sección 4.10, debe cumplirse

$$\mu = (\Delta t)^2 \frac{1}{6} c^3 - c (\Delta x)^2 \frac{1}{6} , \quad (125)$$

porque de este modo se anula el primer término y, siempre que al refinar la malla Δx sea proporcional a Δt , el resto de los términos quedan al menos de orden 3 al respecto de Δt . Adicionalmente, si se sustituyen las expresiones (57) y (71) (o bien (73)):

$$\begin{aligned}
C &= \frac{c \Delta t}{\Delta x} , \\
i_p &= \frac{1}{\Delta x} ,
\end{aligned}$$

en la ecuación (125) se obtiene el valor de μ en términos de C e i_p :

$$\mu = \frac{c}{6i_p^2} (C^2 - 1) .$$

Entonces una manera de expresar la ecuación modificada correspondiente a este esquema numérico es:

$$\tilde{u}_t + c\tilde{u}_x = \left((\Delta t)^2 \frac{1}{6} c^3 - c (\Delta x)^2 \frac{1}{6} \right) \tilde{u}_{xxx} ,$$

que es una ecuación dispersiva y su solución exacta se muestra en la sección 6.2.

6.7. Valor de μ para el esquema de Wendroff.

Por la cantidad de términos que implica el error de truncamiento de este esquema no es cómodo escribirlo en una sola ecuación. En lugar de eso, veamos la ecuación (102) y los desarrollos de sus términos en serie de Taylor expresados en las ecuaciones (103), (104), (105) y (106). Ahora sustituimos u por \tilde{u} solución de la ecuación (115),

lo cual implica que satisface las relaciones (122), y tenemos:

$$\begin{aligned} \frac{1}{2} \left(\frac{\tilde{u}(x + \Delta x, t + \Delta t) - \tilde{u}(x + \Delta x, t)}{\Delta t} \right) &= \frac{1}{2} (-c\tilde{u}_x + \mu\tilde{u}_{xxx}) + \Delta x \frac{1}{2} (-c\tilde{u}_{xx} + \mu\tilde{u}_{xxxx}) \\ &\quad + \Delta t \frac{1}{4} (c^2\tilde{u}_{xx} + \mu(-c\tilde{u}_{xxxx} + \tilde{u}_{xxxxt})) \\ &\quad + (\Delta x)^2 \frac{3}{12} (-c\tilde{u}_{xxx} + \mu\tilde{u}_{xxxxx}) \\ &\quad + (\Delta x)(\Delta t) \frac{3}{12} (c^2\tilde{u}_{xxx} + \mu(-c\tilde{u}_{xxxxx} + \tilde{u}_{xxxxxt})) \\ &\quad + (\Delta t)^2 \frac{1}{12} (-c^3\tilde{u}_{xxx} + \mu(c^2\tilde{u}_{xxxxx} - c\tilde{u}_{xxxxxt} + \tilde{u}_{xxxxxtt})) \\ &\quad + O((\Delta x)^3 + (\Delta t)^3), \end{aligned}$$

$$\begin{aligned} \frac{1}{2} \left(\frac{\tilde{u}(x, t + \Delta t) - \tilde{u}(x, t)}{\Delta t} \right) &= \frac{1}{2} (-c\tilde{u}_x + \mu\tilde{u}_{xxx}) + \Delta t \frac{1}{4} (c^2\tilde{u}_{xx} + \mu(-c\tilde{u}_{xxxx} + \tilde{u}_{xxxxt})) \\ &\quad + (\Delta t)^2 \frac{1}{12} (-c^3\tilde{u}_{xxx} + \mu(c^2\tilde{u}_{xxxxx} - c\tilde{u}_{xxxxxt} + \tilde{u}_{xxxxxtt})) \\ &\quad + O((\Delta x)^3), \end{aligned}$$

$$\begin{aligned} c \frac{1}{2} \left(\frac{\tilde{u}(x + \Delta x, t + \Delta t) - \tilde{u}(x, t + \Delta t)}{\Delta x} \right) &= \frac{c}{2} \tilde{u}_x + \Delta x \frac{c}{4} \tilde{u}_{xx} + \Delta t \frac{c}{2} (-c\tilde{u}_{xx} + \mu\tilde{u}_{xxxx}) \\ &\quad + (\Delta x)^2 \frac{c}{2} \tilde{u}_{xxx} + (\Delta x)(\Delta t) \frac{3c}{12} (-c\tilde{u}_{xxx} + \mu\tilde{u}_{xxxxt}) \\ &\quad + (\Delta t)^2 \frac{3c}{12} (c^2\tilde{u}_{xxx} + \mu(-c\tilde{u}_{xxxxx} + \tilde{u}_{xxxxxt})) + O((\Delta x)^3 + (\Delta t)^3), \end{aligned}$$

$$c \frac{1}{2} \left(\frac{\tilde{u}(x + \Delta x, t) - \tilde{u}(x, t)}{\Delta x} \right) = \frac{c}{2} \tilde{u}_x + \Delta x \frac{c}{4} \tilde{u}_{xx} + (\Delta x)^2 \frac{c}{12} \tilde{u}_{xxx} + O((\Delta x)^3).$$

Al sumar las cuatro expresiones se obtiene

$$L = \tilde{u}_{xxx} \left(\mu + (\Delta x)^2 \frac{c}{12} + (\Delta t)^2 \frac{c}{12} \right) + \mu O(\Delta t + \Delta x) + O((\Delta t)^3 + (\Delta x)^3),$$

de donde se observa que, para que el esquema numérico de Wendroff sea de tercer orden con respecto a su ecuación modificada correspondiente (bajo la definición descrita en la sección 4.10), es necesario que

$$\mu = \frac{c}{12} (\Delta x)^2 - \frac{c^3}{12} (\Delta t)^2. \quad (126)$$

Por lo tanto, la ecuación modificada correspondiente al esquema de Wendroff es

$$\tilde{u}_t + c\tilde{u}_x = \left(\frac{c}{12} (\Delta x)^2 - \frac{c^3}{12} (\Delta t)^2 \right) \tilde{u}_{xxx},$$

y su solución exacta se muestra en la sección 6.2.

Adicionalmente, podemos obtener el valor de μ en términos de C e i_p mediante sustituir las expresiones (57) y

(71) (o bien (73)):

$$C = \frac{c \Delta t}{\Delta x},$$

$$i_p = \frac{1}{\Delta x},$$

en la ecuación (126), el resultado es:

$$\mu = \frac{c}{12i_p^2} (1 - C^2).$$

6.8. Implementación y comparación con resultados numéricos.

Las soluciones de las ecuaciones modificadas, como se ha visto en las secciones 6.1 y 6.2, son expresiones analíticas y podría pensarse que estas no requieren cálculo numérico para ofrecer datos comparables con los resultados expuestos en la sección 5.6. Sin embargo, como dichas soluciones están expresadas como sumas de funciones senoidales es necesario realizar esta suma para cada uno de los puntos que se quiera graficar y esto requiere la discusión de ciertos detalles relacionados con la implementación.

Para implementar esta suma consideraremos que la condición inicial de las soluciones de las ecuaciones modificadas es periódica, con longitud de onda λ_M , incluso en el caso de modelar resultados numéricos obtenidos a partir de una condición inicial de soporte compacto. En el caso de propagaciones numéricas de condiciones iniciales periódicas, el modo en que se han definido este tipo de condiciones iniciales (en la sección 4.12.1) permiten definir $\lambda_M = 1$; y la suma (de funciones senoidales) consta de un solo término porque el perfil de la onda es precisamente senoidal. Por otro lado, al modelar mediante la solución de ecuaciones modificadas los resultados numéricos obtenidos a partir de condiciones iniciales de soporte compacto, lo único que será necesario es tomar un valor de λ_M suficientemente grande como para que la señal, contemplando su propagación y su deformación, es decir, los datos que efectivamente serán graficados, estén dentro del intervalo $[0, \lambda_M]$ y en este caso se ha usado un máximo de 256 modos, es decir, un máximo de 256 términos en la sumatoria, y se ha verificado que usando más términos no se gana en detalle en cuanto a la resolución de las gráficas que se han usado.

Dado que la ecuación modificada correspondiente a los esquemas contracorriente y de Lax-Friedrichs es una ecuación difusiva y no dispersiva, podemos esperar que su solución modele la difusión numérica que se ha observado en estos esquemas pero no así la dispersión numérica que también, suponemos, se ha observado. Del mismo modo, para los esquemas de Lax-Wendroff y Wendroff la ecuación modificada que les corresponde es una ecuación dispersiva y no difusiva, así que esperamos que modele la dispersión numérica observada para estos esquemas pero no la difusión numérica que sí se ha observado para el esquema de Lax-Wendroff.

Los valores obtenidos de α o μ para cada uno de los esquemas numéricos se encuentran en las ecuaciones (123),

(124), (125), (126), y son respectivamente:

$$\begin{aligned} \text{Contracorriente:} \quad & \alpha = \Delta x \frac{c}{2} - \Delta t \frac{1}{2} c^2, \\ \text{Lax-Friedrichs:} \quad & \alpha = \frac{(\Delta x)^2}{2(\Delta t)} - \frac{1}{2} (\Delta t) c^2, \\ \text{Lax-Wendroff:} \quad & \mu = (\Delta t)^2 \frac{1}{6} c^3 - c (\Delta x)^2 \frac{1}{6}, \\ \text{Wendroff:} \quad & \mu = \frac{c}{12} (\Delta x)^2 - \frac{c^3}{12} (\Delta t)^2. \end{aligned}$$

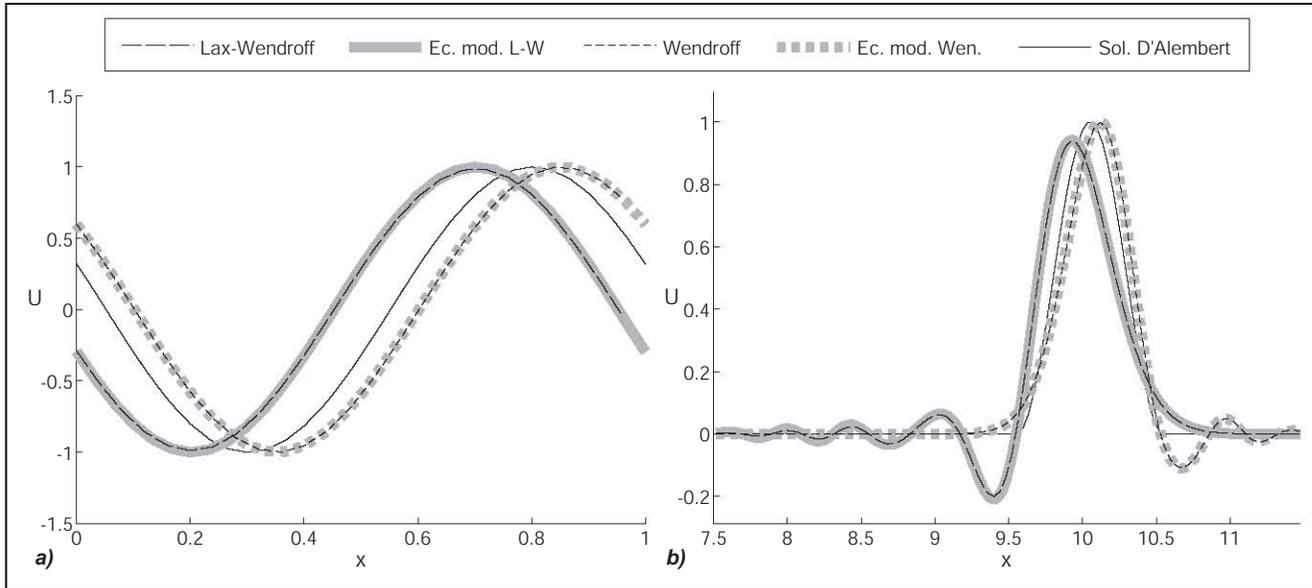


Figura 53: Resultados numéricos y ecuaciones modificadas para los esquemas de Lax-Wendroff y Wendroff, con $i_p = 25$, $C = 0.1$ y $j = 2389$, para condiciones iniciales: **a)** periódicas y **b)** de soporte compacto.

Mediante estas expresiones podemos construir gráficas como las que se muestran en las figuras 53, 54, 55 y 56 que nos permiten observar que ciertamente los resultados numéricos tienen una proximidad mucho mejor con sus respectivas ecuaciones modificadas que con la solución de D'Alembert. Como se ha mencionado, para los primeros dos casos (contracorriente y Lax-Friedrichs) la ecuación modificada es puramente difusiva y para los esquemas de Lax-Wendroff y Wendroff la ecuación modificada que corresponde es puramente dispersiva; esta proximidad nos sugiere que para los dos primeros esquemas es más importante la difusión numérica y para los dos últimos es más importante la dispersión numérica. La correspondencia de las soluciones de las ecuaciones modificadas con los resultados numéricos obtenidos a partir de condiciones iniciales de soporte compacto nos indica hasta qué grado la deformación del perfil es debida a difusión numérica en los dos primeros esquemas y hasta qué grado es debida a dispersión numérica en los dos últimos. En la mayoría de los casos esta correspondencia es bastante buena. Por otro lado, en las figuras 55 y 56 se observa una diferencia entre los resultados numéricos y las soluciones de ecuaciones modificadas. Una de las causas de esta diferencia es la presencia de dispersión numérica en los esquemas de Lax-Friedrichs y contracorriente, así como de difusión numérica en el esquema de Lax-Wendroff. Una discusión más detallada de esta divergencia se encuentra en la sección 6.9.

Ahora buscamos comparar las propiedades de las ecuaciones modificadas con las tendencias de dispersión numérica y difusión numérica de los esquemas de la ecuación de advección, que se han señalado en la sección

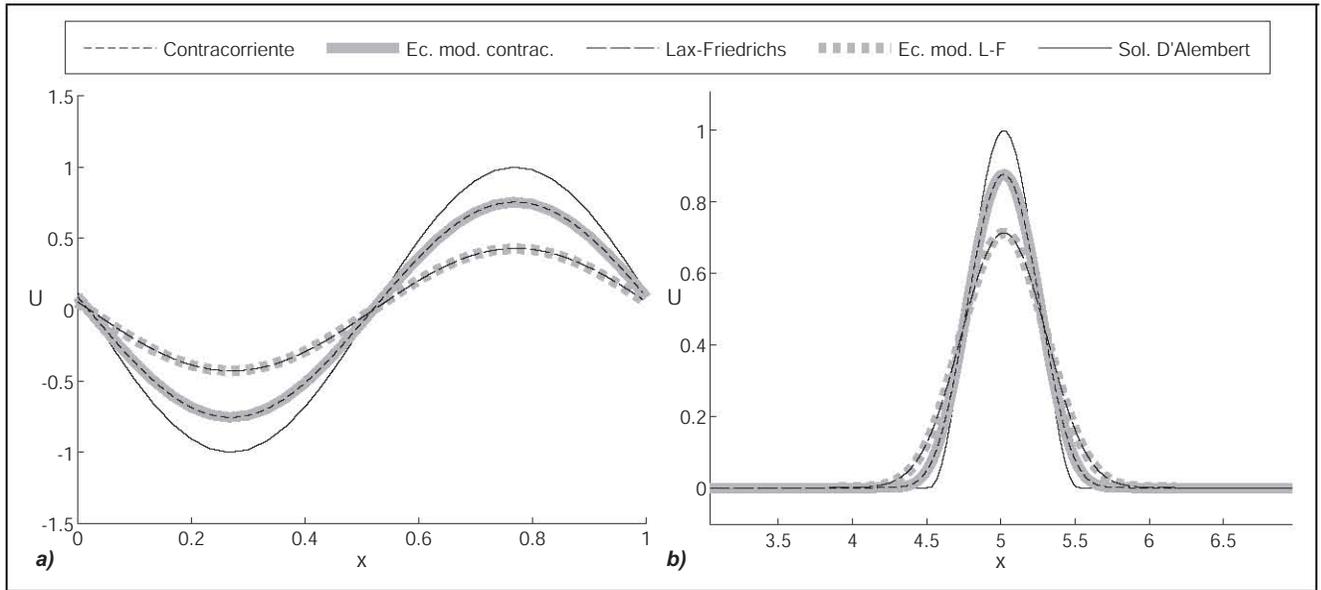


Figura 54: Resultados numéricos y ecuaciones modificadas para los esquemas contracorriente y de Lax-Friedrichs, con $i_p = 158$, $C = 0.5$ y $j = 1429$, para condiciones iniciales: **a)** periódicas y **b)** de soporte compacto.

5.6. Como el análisis de dichos resultados numéricos se ha hecho tomando j como variable independiente y las ecuaciones modificadas han sido resueltas en términos de x y t , es necesario expresar las soluciones de las ecuaciones modificadas como función de los índices i y j . Para ello usamos la definición de la discretización, en la expresión (56), que señala

$$\begin{aligned} x &= x_L + (i - 1) \Delta x, \\ t &= (j - 1) \Delta t. \end{aligned}$$

Entonces al reescribir las soluciones de la ecuaciones modificadas obtenemos para el caso de la solución (117) de la ecuación de advección-difusión (114):

$$\tilde{u} = \sum_{k=0}^{\infty} \left(\exp \left(-\alpha \left(\frac{2\pi k}{\lambda_M} \right)^2 (j - 1) \Delta t \right) \cdot D_k \sin \left(\frac{2\pi k}{\lambda_M} (x_L + (i - 1) \Delta x - c(j - 1) \Delta t) + \phi_k \right) \right),$$

y para el caso de la solución (120) de ecuación de advección con dispersión (115):

$$\tilde{u} = \sum_{k=0}^{\infty} \left(D_k \sin \left(\frac{2\pi}{\lambda_M} k (x_L + (i - 1) \Delta x) - \left(\frac{2\pi}{\lambda_M} ck + \frac{(2\pi)^3}{\lambda_M^3} \mu k^3 \right) (j - 1) \Delta t + \phi_k \right) \right).$$

De este modo, en el primer caso la amplitud de los modos decae exponencialmente al respecto de las variables i , j , de acuerdo a una constante $\tau'_k = \tau_k \Delta t$, donde, como indica la ecuación (116),

$$\tau_k = \alpha \left(\frac{2\pi k}{\lambda_M} \right)^2.$$

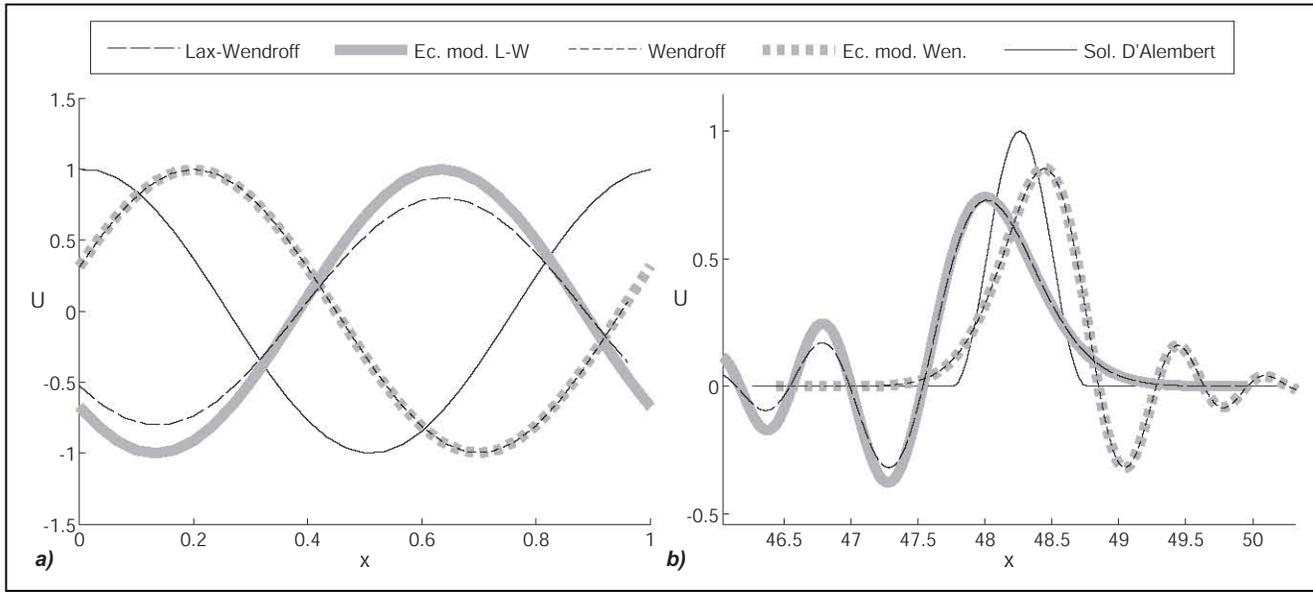


Figura 55: Resultados numéricos y ecuaciones modificadas para los esquemas de Lax-Wendroff y Wendroff, con $i_p = 25$, $C = 0.5$ y $j = 2389$, para condiciones iniciales: **a)** periódicas y **b)** de soporte compacto.

En el segundo caso los modos se propagan con una velocidad

$$\begin{aligned}\tilde{c}'_k &= \frac{\left(\frac{2\pi}{\lambda_M}ck + \frac{(2\pi)^3}{\lambda_M^3}\mu k^3\right)\Delta t}{\frac{2\pi}{\lambda_M}k\Delta x} \\ &= \frac{\Delta t}{\Delta x} \left(c + \mu \frac{(2\pi)^2}{\lambda_M^2}k^2\right) \\ &= \frac{\Delta t}{\Delta x} \tilde{c}_k,\end{aligned}$$

donde, como se ha expuesto en la ecuación (119),

$$\tilde{c}_k = c + \mu \frac{(2\pi)^2}{\lambda_M^2}k^2.$$

Entonces resulta conveniente calcular el valor de τ'_k o \tilde{c}'_k , según sea conveniente, para cada uno de los esquemas numéricos, aprovechando nuevamente las ecuaciones (123), (124), (125), (126) para obtener:

$$\begin{aligned}\text{Contracorriente:} \quad \tau'_k &= \Delta t \left(\frac{2\pi k}{\lambda_M}\right)^2 \left(\Delta x \frac{c}{2} - \Delta t \frac{1}{2}c^2\right), \\ \text{Lax-Friedrichs:} \quad \tau'_k &= \Delta t \left(\frac{2\pi k}{\lambda_M}\right)^2 \left(\frac{(\Delta x)^2}{2(\Delta t)} - \frac{1}{2}(\Delta t)c^2\right), \\ \text{Lax-Wendroff:} \quad \tilde{c}'_k &= \frac{\Delta t}{\Delta x} \left(c + \left((\Delta t)^2 \frac{1}{6}c^3 - c(\Delta x)^2 \frac{1}{6}\right) \frac{(2\pi)^2}{\lambda_M^2}k^2\right), \\ \text{Wendroff:} \quad \tilde{c}'_k &= \frac{\Delta t}{\Delta x} \left(c + \left(\frac{c}{12}(\Delta x)^2 - \frac{c^3}{12}(\Delta t)^2\right) \frac{(2\pi)^2}{\lambda_M^2}k^2\right).\end{aligned}$$

Luego, reescribimos estas ecuaciones en términos de los parámetros C e i_p , que se han usado como base de la

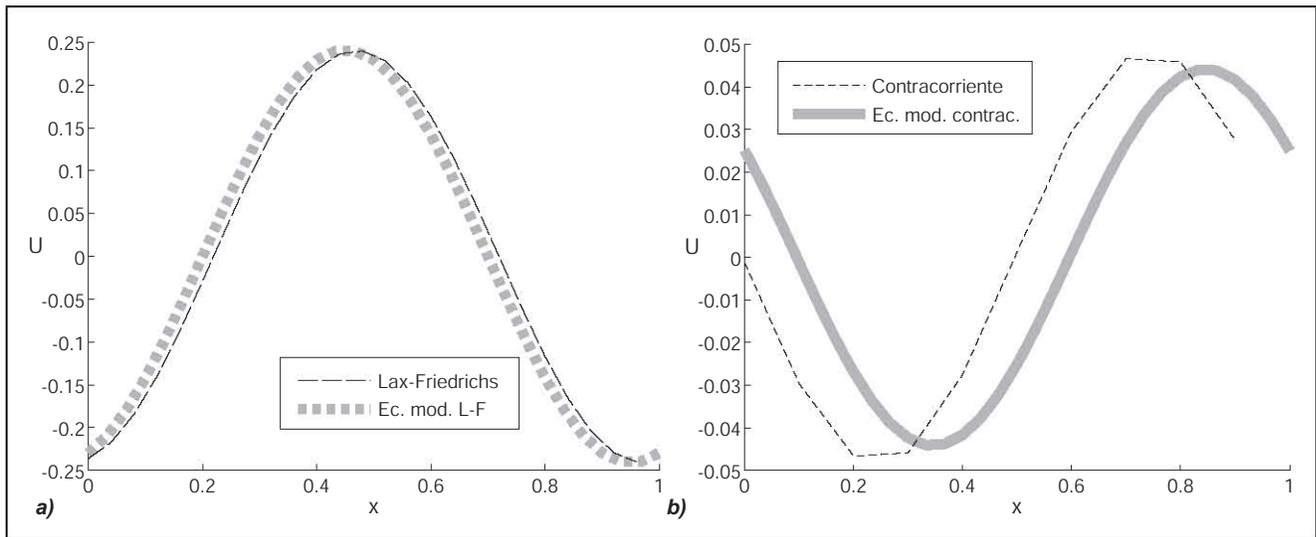


Figura 56: Resultados numéricos y ecuaciones modificadas a partir de condiciones iniciales periódicas para los esquemas: **a)** de Lax-Friedrichs, con $i_p = 25$, $C = 0.5$ y $j = 61$; y **b)** contracorriente, con $i_p = 10$, $C = 0.01$ y $j = 1597$.

descripción de los resultados numéricos en la sección 5.6, y por medio de las identidades (71) (o bien (73)) y (57),

$$\Delta x = \frac{1}{i_p},$$

$$C = \frac{c \Delta t}{\Delta x},$$

obtenemos:

$$\text{Contraorriente:} \quad \tau'_k = \frac{1}{i_p^2} \left(\frac{\pi k}{\lambda_M} \right) (C - C^2), \quad (127)$$

$$\text{Lax-Friedrichs:} \quad \tau'_k = \frac{1}{i_p^2} \left(\frac{\pi k}{\lambda_M} \right) (1 - C^2), \quad (128)$$

$$\text{Lax-Wendroff:} \quad \tilde{c}'_k = C + \frac{1}{6i_p^2} \left(\frac{2\pi}{\lambda_M} k \right)^2 (C^3 - C), \quad (129)$$

$$\text{Wendroff:} \quad \tilde{c}'_k = C + \frac{1}{12i_p^2} \left(\frac{2\pi}{\lambda_M} k \right)^2 (C - C^3). \quad (130)$$

Luego notamos que si escribimos la solución de D'Alembert en las variables i, j (aprovechando nuevamente las ecuaciones (56), (71), (73) y (57)) se obtiene:

$$\begin{aligned} u(x, t) &= f(x - ct) \\ &= f(x_L + (i - 1)\Delta x - c(j - 1)\Delta t) \\ &= f\left(x_L + (i - 1)\frac{1}{i_p} - \frac{C}{i_p}(j - 1)\right), \end{aligned}$$

donde se observa que bajo estas variables el perfil correspondiente a la solución de D'Alembert se desplaza (hacia

la derecha si $C > 0$) con una velocidad igual a

$$-\frac{-\frac{C}{i_p}}{\frac{1}{i_p}} = C,$$

de modo que la parte de \tilde{c}'_k que nos indica la presencia de dispersión es precisamente el segundo término en las expresiones (129) y (130).

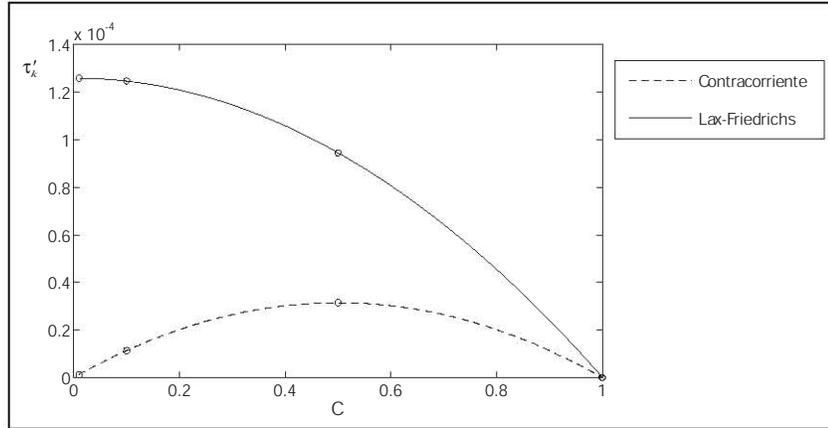


Figura 57: τ'_k como función de C para los esquemas contracorriente y de Lax-Friedrichs, con $i_p = 158$.

Entonces graficamos, para un valor fijo de i_p , en la figura 57 los valores de τ'_k como función de C y en la figura 58 los valores de $\tilde{c}'_k - C$ como función de C . En ambas figuras se han indicado con círculos los valores de C para los cuales se han presentado resultados numéricos en la sección 5.6. No es necesario hacer las gráficas de estos parámetros como función de i_p ya que de las expresiones (127), (128), (129), (130) se observa directamente que todos estos parámetros guardan una relación inversamente proporcional con respecto a i_p^2 .

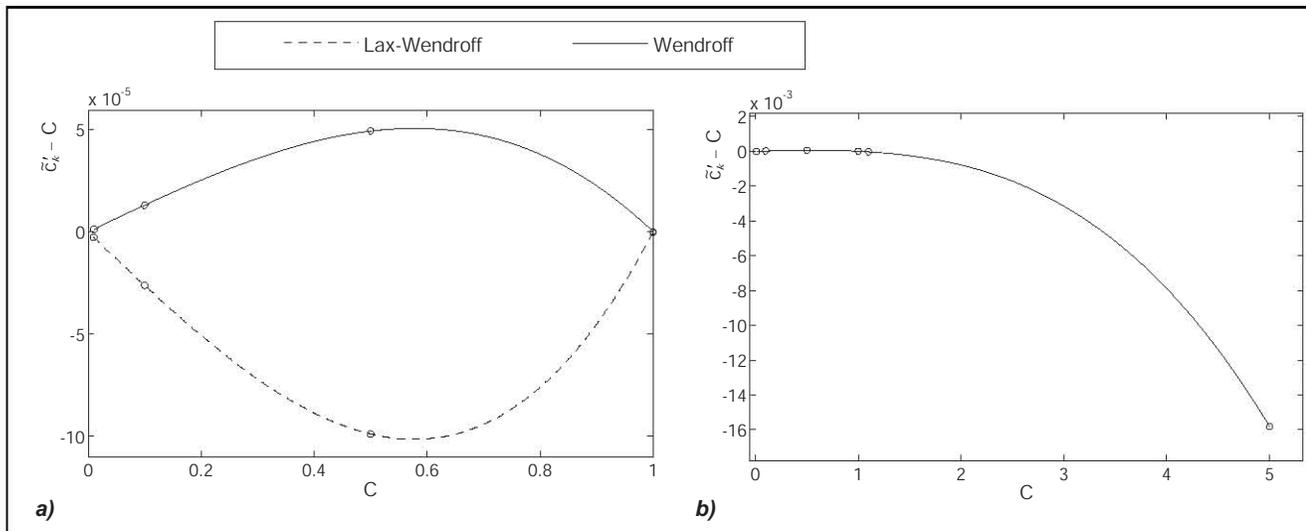


Figura 58: $\tilde{c}'_k - C$ como función de C para los esquemas de Lax-Wendroff y Wendroff, con $i_p = 158$. La figura a es un detalle de la figura b.

Estos resultados confirman las tendencias señaladas en la sección 5.6 para difusión numérica en los esquemas contracorriente y de Lax-Friedrichs: de mayor magnitud para el esquema de Lax-Friedrichs; de menor magnitud en

todos los casos para i_p mayor; también de menor magnitud para C cercano a 1 en el esquema de Lax-Friedrichs; y en cuanto al esquema contracorriente, de menor magnitud para $C = 0.1$ que para $C = 0.5$, y de magnitud aún menor para $C = 0.01$. Del mismo modo confirman las tendencias señaladas para dispersión numérica de los esquemas Lax-Wendroff y Wendroff: de mayor magnitud y en sentido negativo para el esquema de Lax-Wendroff; en sentido positivo para $C < 1$ y en sentido opuesto para $C > 1$ en el esquema de Wendroff; y de mayor magnitud, en el rango $C < 1$, para el esquema de Lax-Wendroff. Finalmente las figuras 57 y 58 confirman que la magnitud de la dispersión y difusión numérica no guarda una relación monótona con respecto a C ; para conocer dicha relación sería necesario realizar propagaciones numéricas con más valores para C , en particular dentro del intervalo $(0.5, 1)$.

Como se ha dicho antes, por la naturaleza de la ecuaciones modificadas, no podemos esperar que éstas describan la dispersión numérica presente en los esquemas contracorriente y de Lax-Friedrichs, ni la difusión numérica presente en el esquema de Lax-Wendroff.

6.9. Otras ecuaciones modificadas.

Como se ha mencionado en las secciones anteriores, para algunos casos es posible observar que existe una discrepancia entre la ecuación modificada de cada esquema y los resultados numéricos que le corresponden. Es notable que esta discrepancia, en el caso de contracorriente y Lax-Friedrichs, incluye un corrimiento lateral, y en el caso de Lax-Wendroff incluye un cambio en la amplitud.

Si miramos el modo como hemos construido las ecuaciones modificadas, se observa que siempre debe existir una ecuación diferencial cuya solución \tilde{u} , al ser sustituida en la expresión que describe el error de truncamiento local L de un determinado esquema numérico, elimina tantos términos como se quiera; es decir, siempre existe una ecuación diferencial con respecto a la cual el esquema numérico es de orden n , para cualquier $n \in \mathbb{Z}^+$. Dicha ecuación diferencial se encuentra implícitamente en la expresión que define al error de truncamiento local de cada esquema numérico.

Así pues, como la aproximación de los resultados numéricos a estas nuevas ecuaciones modificadas debe ser mejor, puede esperarse que dichas ecuaciones modificadas sean dispersivas para los esquemas contracorriente y de Lax-Friedrichs; y sean difusivas para el esquema de Lax-Wendroff. Sin embargo, tenemos una ventaja en cuanto a las ecuaciones modificadas que se han descrito en secciones anteriores, y ésta es que podemos expresar la solución exacta de las ecuaciones modificadas explícitamente a partir de funciones elementales, mientras que para otras ecuaciones modificadas no tenemos elementos que nos garanticen que también pudiéramos conocer su solución explícitamente.

7. Esquemas numéricos relativos a la ecuación de Burgers no viscosa.

Al hacer las propagaciones numéricas usando los esquemas respectivos a la ecuación de advección, se han advertido algunos aspectos que indican cuáles de ellos y bajo qué condiciones brindan mejores resultados. Aprovecharemos esta información para hacer propagaciones numéricas que aproximen la solución de la ecuación de Burgers no viscosa del mejor modo que nos sea posible.

La ecuación de Burgers no viscosa tiene una forma similar a la ecuación de advección, en el sentido de que simplemente cambia una constante c por una aparición de u . Esto nos sugiere que debería ser sencillo adecuar los esquemas numéricos que se han usado para aproximar la solución de la ecuación de advección para que aproximen la solución de la ecuación de Burgers no viscosa. Una manera de hacer esta adecuación es simplemente cambiar, en los esquemas numéricos relativos a la ecuación de advección, las apariciones de c por U_i^j y lo haremos de esta manera con los dos esquemas que mejores resultados han ofrecido (en el caso de la ecuación de advección), y que son el esquema de Lax-Wendroff y el de Wendroff. Los esquemas numéricos que se han denominado *primero* y *segundo*, respectivamente, son el resultado de hacer esta sustitución.

Como será mostrado en las secciones 7.1 y 7.2 el resultado de la sustitución mencionada da como resultado dos esquemas numéricos relativos a la ecuación de Burgers no viscosa que son de primer orden. Uno de los aspectos que resalta de los resultados expuestos en la sección 5.6 es que los esquemas de segundo orden brindan resultados notablemente mejores que los de primer orden. Por lo tanto, se ha seguido un procedimiento análogo al expuesto en la sección 5.3 para construir un esquema numérico de segundo orden con respecto a la ecuación de Burgers no viscosa; dicho esquema se ha denominado *tercero* y sus detalles se encuentran en la sección 7.3.

Como se ha hecho notar en la sección 4.6, para los esquemas numéricos relativos a la ecuación de Burgers no podemos utilizar notación polinomial debido a su carácter no lineal. Esto implica que no podemos usar tampoco la notación matricial, y lo que es más importante, no podemos hacer un análisis de estabilidad mediante la condición de Von Neumann bajo el procedimiento que se ha usado para los esquemas relativos a la ecuación de advección, de modo que en el caso de la ecuación de Burgers no viscosa no nos será útil el teorema de equivalencia de Lax y tendremos que hacer las propagaciones numéricas sin saber si los esquemas numéricos son efectivamente convergentes.

7.1. Esquema primero.

Construimos este esquema, que denominamos *primero*, a partir del esquema de Lax-Wendroff, descrito en la sección 5.3. Siguiendo el procedimiento descrito al principio de esta sección 7, en la ecuación (94)

$$U_i^{j+1} = U_i^j - c(\Delta t) \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} + c^2(\Delta t)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{2(\Delta x)^2},$$

sustituimos las apariciones c por U_i^j y se obtiene la ecuación que define a este esquema numérico *primero*:

$$U_i^{j+1} = U_i^j - U_i^j(\Delta t) \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} + (U_i^j \Delta t)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{2(\Delta x)^2}. \quad (131)$$

Como puede observarse en la expresión (131), este es un esquema de tipo explícito.

Aunque, como se ha mencionado antes, no es posible hacer un análisis de estabilidad bajo la prueba de Von Neu-

mann para este esquema numérico, lo que sí es posible y realizaremos a continuación es la evaluación de consistencia y orden. Para evaluar el error de truncamiento L reordenamos los términos de la ecuación (131) en la forma

$$0 = \frac{U_i^{j+1} - U_i^j}{\Delta t} + U_i^j \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} - (U_i^j)^2 \frac{(\Delta t)}{2} \cdot \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2},$$

y luego sustituimos U por u la solución de la ecuación de Burgers no viscosa, evaluada en el punto que le corresponda, del modo que se ha descrito en la sección 4.10, para obtener

$$L = \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} + u(x, t) \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2(\Delta x)} - (u(x, t))^2 \frac{(\Delta t)}{2} \cdot \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2}.$$

Posteriormente observamos el desarrollo de Taylor para cada uno de los términos:

$$\begin{aligned} \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} &= u_t + \Delta t \frac{1}{2} u_{tt} + O((\Delta t)^2), \\ \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2(\Delta x)} &= u_x + O((\Delta x)^2), \\ \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2} &= u_{xx} + O((\Delta x)^2), \end{aligned}$$

por lo que

$$\begin{aligned} L &= u_t + \Delta t \frac{1}{2} u_{tt} + O((\Delta t)^2) + u(u_x + O((\Delta x)^2)) - u^2 \frac{(\Delta t)}{2} (u_{xx} + O((\Delta x)^2)) \\ &= u_t + \Delta t \frac{1}{2} u_{tt} + uu_x - u^2 u_{xx} \frac{(\Delta t)}{2} + (\Delta t) O((\Delta x)^2) + O((\Delta x)^2) + O((\Delta t)^2). \end{aligned}$$

Ahora, como u es solución de la ecuación de Burgers no viscosa, podemos aprovechar las identidades (63) descritas en la sección 4.9, y se obtiene

$$\begin{aligned} L &= -uu_x + \Delta t \frac{1}{2} (u_x^2 + uu_x^2 + uu_{xx}) + uu_x - u^2 u_{xx} \frac{(\Delta t)}{2} + (\Delta t) O((\Delta x)^2) + O((\Delta x)^2) + O((\Delta t)^2) \\ &= \frac{\Delta t}{2} (u_x^2 + uu_x^2 + uu_{xx} - u^2 u_{xx}) + (\Delta t) O((\Delta x)^2) + O((\Delta x)^2) + O((\Delta t)^2) \\ &= O(\Delta t) + (\Delta t) O((\Delta x)^2) + O((\Delta x)^2), \end{aligned}$$

donde puede observarse que, dado que L tiende a cero cuando se refina la malla, este esquema es consistente con la ecuación de Burgers no viscosa. Para encontrar su orden consideramos a la razón $\frac{\Delta t}{\Delta x}$ como una constante y de este modo podemos reducir L a términos de Δt ,

$$\begin{aligned} L &= O(\Delta t) + (\Delta t) O((\Delta t)^2) + O((\Delta t)^2) \\ &= O(\Delta t), \end{aligned}$$

por lo que este esquema *primero* es de primer orden con respecto a la ecuación de Burgers no viscosa.

7.2. Esquema segundo.

De modo análogo al esquema anterior, construiremos este esquema *segundo* para la ecuación de Burgers no viscosa sustituyendo las apariciones de c por U_i^j , pero en este caso a partir del esquema de Wendroff, descrito en la sección 5.4. Entonces tenemos la ecuación (99)

$$\frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_i^{j+1} - U_i^j}{\Delta t} \right) + c \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right) = 0,$$

donde al sustituir c por U_i^j se obtiene

$$\frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_{i+1}^j}{\Delta t} + \frac{U_i^{j+1} - U_i^j}{\Delta t} \right) + U_i^j \frac{1}{2} \left(\frac{U_{i+1}^{j+1} - U_i^{j+1}}{\Delta x} + \frac{U_{i+1}^j - U_i^j}{\Delta x} \right) = 0, \quad (132)$$

o equivalentemente,

$$U_{i+1}^{j+1} = \frac{1 - \frac{\Delta t}{\Delta x} U_i^j}{1 + \frac{\Delta t}{\Delta x} U_i^j} (U_{i+1}^j - U_i^{j+1}) + U_i^j. \quad (133)$$

Como en estas expresiones aparecen tanto U_i^{j+1} como U_{i+1}^{j+1} , este esquema numérico *segundo* es implícito, al igual que el esquema de Wendroff que se ha usado como su base.

El siguiente paso es verificar que este esquema *segundo* sea consistente con la ecuación de Burgers no viscosa. Para ello expresamos su correspondiente error de truncamiento local L , sustituyendo U por u , la solución de la ecuación de Burgers no viscosa en la expresión (132), y obtenemos

$$L = \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x + \Delta x, t)}{\Delta t} + \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \right) + u(x, t) \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x, t + \Delta t)}{\Delta x} + \frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} \right).$$

Luego, calculamos por separado las series de Taylor correspondientes a cada uno de los términos:

$$\frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x + \Delta x, t)}{\Delta t} \right) = \frac{1}{2} u_t + \frac{1}{2} (\Delta x) u_{xt} + \frac{1}{4} (\Delta t) u_{tt} + O((\Delta x)^2 + (\Delta t)^2),$$

$$\frac{1}{2} \left(\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \right) = \frac{1}{2} u_t + \frac{1}{4} (\Delta t) u_{tt} + O((\Delta t)^2),$$

$$u(x, t) \frac{1}{2} \left(\frac{u(x + \Delta x, t + \Delta t) - u(x, t + \Delta t)}{\Delta x} \right) = \frac{1}{2} uu_x + \frac{1}{4} (\Delta x) uu_{xx} + \frac{1}{2} (\Delta t) uu_{xt} + O((\Delta x)^2 + (\Delta t)^2),$$

$$u(x, t) \frac{1}{2} \left(\frac{u(x + \Delta x, t) - u(x, t)}{\Delta x} \right) = \frac{1}{2} uu_x + \frac{1}{4} (\Delta x) uu_{xx} + O((\Delta x)^2),$$

por lo que

$$\begin{aligned}
L &= \frac{1}{2}u_t + \frac{1}{2}(\Delta x)u_{xt} + \frac{1}{4}(\Delta t)u_{tt} + O((\Delta x)^2 + (\Delta t)^2) + \frac{1}{2}u_t + \frac{1}{4}(\Delta t)u_{tt} + O((\Delta t)^2) \\
&\quad + \frac{1}{2}uu_x + \frac{1}{4}(\Delta x)uu_{xx} + \frac{1}{2}(\Delta t)uu_{xt} + O((\Delta x)^2 + (\Delta t)^2) + \frac{1}{2}uu_x + \frac{1}{4}(\Delta x)uu_{xx} + O((\Delta x)^2) \\
&= u_t + \frac{1}{2}(\Delta x)u_{xt} + \frac{1}{2}(\Delta t)u_{tt} + uu_x + \frac{1}{2}(\Delta x)uu_{xx} + \frac{1}{2}(\Delta t)uu_{xt} \\
&\quad + O((\Delta x)^2) + O((\Delta t)^2) + O((\Delta x)^2 + (\Delta t)^2) .
\end{aligned}$$

Ahora, gracias a que u satisface la ecuación de Burgers no viscosa, aprovechamos las identidades (63), descritas en la sección 4.9, para sustituir las derivadas de u que involucren a t por expresiones donde solamente aparezcan derivadas con respecto a x , y se obtiene

$$\begin{aligned}
L &= -uu_x + \frac{1}{2}(\Delta x)(-u_x^2 - uu_{xx}) + \frac{1}{2}(\Delta t)(u_x^2 + uu_x^2 + uu_{xx}) + uu_x + \frac{1}{2}(\Delta x)uu_{xx} + \frac{1}{2}(\Delta t)u(-u_x^2 - uu_{xx}) \\
&\quad + O((\Delta x)^2) + O((\Delta t)^2) + O((\Delta x)^2 + (\Delta t)^2) \\
&= -\frac{1}{2}(\Delta x)u_x^2 + \frac{1}{2}(\Delta t)uu_x^2 + O((\Delta x)^2) + O((\Delta t)^2) + O((\Delta x)^2 + (\Delta t)^2) \\
&= O(\Delta x) + O(\Delta t) + O((\Delta x)^2 + (\Delta t)^2) .
\end{aligned}$$

En esta última expresión se observa que al refinar la malla efectivamente el error de truncamiento tiende a cero; por lo tanto, el esquema *segundo* es consistente con la ecuación de Burgers no viscosa. Por otra parte, si el refinamiento de la malla se hace de modo que $\frac{\Delta t}{\Delta x}$ permanezca constante, entonces Δx es un múltiplo de Δt y

$$\begin{aligned}
L &= O(\Delta t) + O(\Delta t) + O((\Delta t)^2) \\
&= O(\Delta t) ,
\end{aligned}$$

por lo que este esquema es de primer orden.

7.3. Esquema tercero.

El esquema *tercero* se construye de manera diferente a los dos anteriores. Buscando que sea de segundo orden se sigue un procedimiento análogo al descrito en la sección 5.3, en aquel caso con respecto a la ecuación de advección y en este caso con respecto a la ecuación de Burgers no viscosa. Partimos de la serie de Taylor de $u(x, t + \Delta t)$,

$$u(x, t + \Delta t) = u + (\Delta t)u_t + (\Delta t)^2 \frac{1}{2}u_{tt} + O((\Delta t)^3) ,$$

para después utilizar las identidades (63) y obtener

$$\begin{aligned}
u(x, t + \Delta t) &= u - (\Delta t)uu_x + (\Delta t)^2 \frac{1}{2}(2uu_x^2 + u^2u_{xx}) + O((\Delta t)^3) \\
&= u - (\Delta t)uu_x + (\Delta t)^2 uu_x^2 + \frac{(\Delta t)^2}{2}u^2u_{xx} + O((\Delta t)^3) .
\end{aligned}$$

Truncando la serie a segundo orden, sustituimos las derivadas por aproximaciones también a segundo orden (como se han expuesto en la sección 4.1), y sustituimos el resto de las apariciones de u por U con los índices adecuados, con lo que se obtiene

$$U_i^{j+1} = U_i^j - (\Delta t) U_i^j \frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} + (\Delta t)^2 U_i^j \left(\frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} \right)^2 + \frac{(\Delta t)^2}{2} (U_i^j)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2}. \quad (134)$$

La ecuación (134) define al esquema numérico *tercero*, e indica que dicho esquema es de tipo explícito. Lo que debemos hacer ahora es verificar que este esquema numérico sea consistente con la ecuación de Burgers no viscosa, para lo que reescribimos la ecuación (134) en una forma que nos permita evaluar el error de truncamiento L :

$$0 = \frac{U_i^{j+1} - U_i^j}{\Delta t} + U_i^j \frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} - (\Delta t) U_i^j \left(\frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} \right)^2 - \frac{(\Delta t)}{2} (U_i^j)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2}.$$

Luego, del modo habitual, sustituimos las apariciones de U por u la solución de la ecuación de Burgers no viscosa evaluada en el punto que le corresponda, y se obtiene

$$\begin{aligned} L = & \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} \\ & + u(x, t) \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2\Delta x} \\ & + (\Delta t) u(x, t) \left(\frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2\Delta x} \right)^2 \\ & - \frac{(\Delta t)}{2} (u(x, t))^2 \left(\frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2} \right). \end{aligned}$$

Posteriormente hacemos los desarrollos de Taylor para cada uno de los términos

$$\begin{aligned} \frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} &= u_t + \frac{1}{2} (\Delta t) u_{tt} + O((\Delta t)^2) \\ &= -uu_x + \frac{(\Delta t)}{2} (2uu_x^2 + u^2 u_{xx}) \\ &\quad + O((\Delta t)^2), \end{aligned} \quad (135)$$

$$u(x, t) \frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2\Delta x} = uu_x + O((\Delta x)^2), \quad (136)$$

$$(\Delta t) u(x, t) \left(\frac{u(x + \Delta x, t) - u(x - \Delta x, t)}{2\Delta x} \right)^2 = (\Delta t) uu_x^2 + O((\Delta x)^2), \quad (137)$$

$$\frac{(\Delta t)}{2} (u(x, t))^2 \left(\frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2} \right) = \frac{(\Delta t)}{2} u^2 u_{xx} + O((\Delta x)^2), \quad (138)$$

donde en la expresión (135) se ha aprovechado el hecho de que u es solución de la ecuación de Burgers no viscosa

para usar las identidades (63). Al hacer la suma de las expresiones (135), (136), (137) y (138) se tiene

$$\begin{aligned} L &= -uu_x + \frac{(\Delta t)}{2} (2uu_x^2 + u^2u_{xx}) + O((\Delta t)^2) + uu_x + O((\Delta x)^2) \\ &\quad + (\Delta t) uu_x^2 + O((\Delta x)^2) + \frac{(\Delta t)}{2} u^2u_{xx} + O((\Delta x)^2) \\ &= O((\Delta x)^2) + O((\Delta t)^2), \end{aligned}$$

de modo que, dado que L tiende a cero siempre que se refine la malla, el esquema es efectivamente consistente con la ecuación de Burgers no viscosa y además es de segundo orden porque en el caso de que $\frac{\Delta t}{\Delta x}$ sea constante, es posible expresar

$$L = O((\Delta t)^2).$$

7.4. Implementación

Con respecto a los parámetros de control también se han escogido, para la implementación de los esquemas relativos a la ecuación de Burgers no viscosa, solamente los que en el caso de la ecuación de advección han resultado más favorables. La razón de que no se haya hecho un análisis más detallado es que, como será mostrado en la sección 7.5, las propagaciones numéricas realizadas ya permiten observar las limitaciones del tipo de esquemas que se han propuesto para la ecuación de Burgers no viscosa.

En el caso de i_p los resultados expuestos en la sección 5.6 muestran que valores grandes deben traer mejores resultados, por lo tanto se ha escogido $i_p = 1000$, lo cual además permite hacer una comparación con los resultados obtenidos (para el mismo valor de i_p) de los esquemas relativos a la ecuación de advección. Por otra parte, dado que la constante c no aparece en los esquemas de la ecuación de Burgers no viscosa, no es posible reducir las expresiones algebraicas de dichos esquemas en términos del parámetro $C = c \frac{\Delta t}{\Delta x}$, pero es posible hacer una interpretación de C que resulta útil: en el caso de los esquemas de la ecuación de advección el valor de c siempre se ha tomado igual a 1, de modo que el valor de C ha determinado directamente el valor del cociente $\frac{\Delta t}{\Delta x}$, por lo que tomaremos para los esquemas de la ecuación de Burgers no viscosa el valor de dicho cociente que haya resultado más favorable para la ecuación de advección. Este valor corresponde a $C = 1$ porque, como se ha señalado en la sección 5.6, dicho valor implica que los resultados ofrecidos por los esquemas numéricos relativos a la ecuación de advección corresponden a la solución exacta de dicha ecuación. Sin embargo, es notable que en el caso de la ecuación de Burgers no viscosa, establecer $\frac{\Delta t}{\Delta x} = 1$ no trae consigo una simplificación de los esquemas numéricos (como ha sucedido para los esquemas de la ecuación de advección), ni se tienen elementos que señalen que los resultados numéricos de los esquemas de la ecuación de Burgers no viscosa necesariamente correspondan a su solución exacta.

Un aspecto que ha cambiado, con respecto al caso de la ecuación de advección, es que para las propagaciones numéricas respectivas a la ecuación de Burgers no viscosa fue necesario aumentar el número de pasos temporales calculados de 2500 a 7500, porque de esta manera, para los valores escogidos de i_p y $\frac{\Delta t}{\Delta x}$, pueden observarse la formación de discontinuidades y el comportamiento de los esquemas numéricos, que serán descritos en la sección 7.5.

Del mismo modo que se ha hecho con la ecuación de advección, se han utilizado para los esquemas relativos a la ecuación de Burgers no viscosa dos tipos de condiciones iniciales: periódicas y de soporte compacto, como se han descrito en la sección 4.12, salvo que en este caso se ha usado un nivel de equilibrio igual a 0.5 y una amplitud

igual a 0.1. Nuevamente los detalles de la implementación con respecto a los esquemas explícitos e implícito son ligeramente distintos y se presentan en las siguientes secciones. La elección hecha para i_p y $\frac{\Delta t}{\Delta x}$ implícitamente fija los valores de Δx y Δt , debido a que en la sección 4.12 se ha establecido para las condiciones iniciales que longitud de onda, o bien la longitud del soporte, es igual a 1, entonces $\Delta x = \Delta t = 0.001$.

7.4.1. Esquemas explícitos: primero y tercero.

Las ecuaciones que se han usado para la implementación de estos esquemas son (131) y (134)

$$\text{Primero: } U_i^{j+1} = U_i^j - U_i^j (\Delta t) \frac{U_{i+1}^j - U_{i-1}^j}{2(\Delta x)} + (U_i^j \Delta t)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{2(\Delta x)^2},$$

$$\text{Tercero: } U_i^{j+1} = U_i^j - (\Delta t) U_i^j \frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} + (\Delta t)^2 U_i^j \left(\frac{U_{i+1}^j - U_{i-1}^j}{2\Delta x} \right)^2 + \frac{(\Delta t)^2}{2} (U_i^j)^2 \frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{(\Delta x)^2}.$$

Como se ha señalado en la sección 4, se usan las ecuaciones mencionadas sobre los valores de la condición inicial para calcular el valor de U correspondiente a sus primeros vecinos, y luego con los valores obtenidos se calculan los correspondientes al siguiente nivel temporal, y así sucesivamente. Los únicos casos que requieren atención especial son, bajo condiciones iniciales periódicas, los extremos del intervalo $i = 1$ y $i = i_p$. En estos casos, al igual que para los esquemas explícitos de la ecuación de advección, se usan los puntos de extremos opuestos; por ejemplo, para calcular el valor de $U_{i_p}^{j+1}$, en el lugar que debería usarse $U_{i_p+1}^j$ se usa U_1^j .

7.4.2. Esquema implícito: segundo.

La ecuación que se ha usado para implementar este esquema es la ecuación (133)

$$U_{i+1}^{j+1} = \frac{1 - \frac{\Delta t}{\Delta x} U_i^j}{1 + \frac{\Delta t}{\Delta x} U_i^j} (U_{i+1}^j - U_i^{j+1}) + U_i^j.$$

La dificultad adicional de esta implementación, con respecto a las que se han expuesto en la sección anterior, es que para calcular el valor de U_{i+1}^{j+1} es necesario conocer U_i^{j+1} , que se encuentra en el mismo nivel temporal.

Para el caso de condiciones iniciales de soporte compacto el problema no es grave, es posible seguir un procedimiento totalmente análogo al caso expuesto en la sección 5.5.3: la información en el esquema segundo fluye de izquierda a derecha, en el sentido de que utiliza el valor de U en puntos a la izquierda para calcular el valor de U de puntos a la derecha, entonces para calcular el valor de U_1^{j+1} por medio de la ecuación (133) se utiliza un cero en el lugar de U_i^{j+1} . Este procedimiento no añade ninguna hipótesis o aproximación, porque suficientemente lejos del soporte de la condición inicial U vale cero para cualquier valor de j y al aplicar la ecuación (133) sobre dichos puntos, se tiene que todos los valores de U a la izquierda de U_1^{j+1} son precisamente cero para toda j .

También en el caso de condiciones iniciales periódicas puede hacerse un desarrollo análogo al expuesto en la sección 5.5.3, esto es, una sustitución recursiva de la ecuación (133) en sí misma (pero evaluada en un punto que quede a un costado). Dicho proceso recursivo finaliza cuando se obtiene de vuelta el punto con el cual se ha comenzado, y esto permite despejar dicho valor en términos de valores que están todos en el nivel temporal anterior.

Si se parte del punto correspondiente a $i = 1$, el resultado de dicho procedimiento es:

$$U_1^{j+1} = \left(1 + (-1)^{i_p-1} \beta_{i_p-1}\right)^{-1} \left(U_{i_p}^j + (-1)^{i_p-1} \beta_{i_p-1} U_2^j + \gamma_{i_p} (U_1^j - U_{i_p-1}^j) + \sum_{s=1}^{i_p-1} (-1)^i \beta_s (U_{i_p+1-s}^j - U_{i_p-1-s}^j) \right), \quad (139)$$

donde,

$$\beta_i = \prod_{s=0}^i \gamma_{i_p-s},$$

$$\gamma_i = \frac{1 - U_i^j \frac{\Delta t}{\Delta x}}{1 + U_i^j \frac{\Delta t}{\Delta x}}.$$

De este modo, en cada nivel temporal, una vez conocido el valor de U_1^j por medio de la ecuación (139), puede obtenerse el valor de U_2^j por medio de la ecuación (133), luego por medio de la misma ecuación el valor de U_3^j , y así sucesivamente hasta llegar a $U_{i_p}^j$.

7.5. Resultados

A continuación se presentan algunas gráficas donde pueden observarse los resultados numéricos obtenidos usando los tres esquemas que se han descrito en las secciones anteriores. Se ha graficado también una serie correspondiente a la solución exacta de la ecuación de Burgers no viscosa construida por características resolviendo la onda de choque por medio de la ley de áreas iguales de Whitham [36], esta solución se ha expuesto en la sección 3.1. Una gama más amplia de resultados puede consultarse en el material digital interactivo cuyo uso se encuentra descrito en el apéndice A.

En las gráficas 59 y 63, correspondientes a $j = 830$ (como en todos los casos $\Delta t = 0.001$, también correspondientes a $t = 0.829$ porque la condición inicial corresponde a $t = 0$), puede observarse que mientras el perfil no presente discontinuidades los esquemas se comportan de modo muy similar a lo que se ha observado para la ecuación de advección. Entre los esquemas de primer orden, el esquema implícito (esquema segundo) cuenta con mayor exactitud que el esquema primero, y por otro lado, el esquema a segundo orden (esquema tercero) ofrece un resultado mejor que los dos anteriores. Es notable que en estas condiciones los tres esquemas aproximan bastante bien a la solución de la ecuación de Burgers no viscosa. De hecho, es necesario hacer un acercamiento muy pronunciado para apreciar la diferencia entre las propagaciones numéricas y la solución analítica, como se ve en 59b y 63b.

A partir de la solución analítica de la ecuación de Burgers no viscosa es posible determinar que la discontinuidad se forma, para el caso de condiciones iniciales periódicas precisamente en $j = 1593$ ($t = 1.592$), y para el caso de condiciones iniciales de soporte compacto en $j = 3181$ ($t = 3.180$). Es posible observar en las figuras 60 y 64 que para un tiempo anterior, pero muy cercano a la formación de choque ($j = 1546$, $t = 1.545$ para condiciones iniciales periódicas y $j = 3016$, $t = 3.015$ para condiciones iniciales de soporte compacto, respectivamente), el comportamiento de los tres esquemas numéricos difiere muy poco de la solución exacta. Algunos pasos temporales más adelante, como lo muestran la figura 61 (para condiciones iniciales periódicas con $j = 1697$, $t = 1.696$) y la figura 65 (para condiciones iniciales de soporte compacto con $j = 3393$, $t = 3.392$), es posible distinguir que los resultados numéricos describen un perfil que no corresponde a la solución analítica. Los detalles de estas figuras

(61b y 65b) muestran que en los resultados numéricos surgen, a partir de la zona donde se presenta la discontinuidad, oscilaciones hacia la derecha en el caso de el esquema segundo y hacia la izquierda para los otros dos esquemas. Luego, en las figuras 62 y 66 se observa que las oscilaciones surgidas a partir de la discontinuidad deterioran la calidad de la aproximación numérica de dos maneras diferentes: para los casos de los esquemas primero y tercero el punto donde aparece la discontinuidad (de los esquemas numéricos) se va alejando del punto donde aparece la discontinuidad de la solución analítica; y para el caso del esquema segundo la oscilación se propaga deformando completamente el perfil.

A partir de lo anterior se concluye que para todos estos esquemas numéricos, a pesar de que se han usado los parámetros que en el caso de la ecuación de advección han brindado mejores resultados, los resultados obtenidos no son buenas aproximaciones a la solución de la ecuación de Burgers no viscosa para tiempos posteriores a la formación de choques, en el sentido de que: en la zona donde se presenta la discontinuidad la calidad de la aproximación disminuye de modo muy sustancial¹¹. Atribuimos este hecho a la manera como hemos construido los esquemas numéricos, aproximando las derivadas de u por medio de discretizaciones algebraicas. A grandes rasgos, las aproximaciones que se hacen en los esquemas de diferencias finitas a los valores de las derivadas parciales, son pendientes de rectas que pasan sobre puntos del tipo (x, t, U_i^j) (donde los valores i, j deben ser tales que correspondan a los valores apropiados de x, t). Dichas pendientes están acotadas superiormente porque, para que el esquema aproxime a la solución analítica, su amplitud debe estar acotada (es decir, el esquema debe ser estable), y además la distancia entre los puntos de la malla es pequeña pero finita. Ahora bien, una discontinuidad como la que se presenta en la solución de la ecuación de Burgers no viscosa corresponde a un punto donde la pendiente es infinita. Por lo tanto, resulta consistente el hecho de que un esquema construido aproximando derivadas por medio de diferencias finitas falle precisamente en este punto. Concluimos, pues, que los métodos de diferencias finitas, en su conjunto, fallan a partir del momento en que las pendientes, afectadas por la no linealidad de la ecuación, se vuelven lo suficientemente pronunciadas como para que ya no se les pueda aproximar de manera discreta.

¹¹Este hecho es importante porque en la mayor parte de los casos, aunque ciertamente no en todos, la zona donde se desea tener mayor resolución es precisamente al rededor de la discontinuidad.

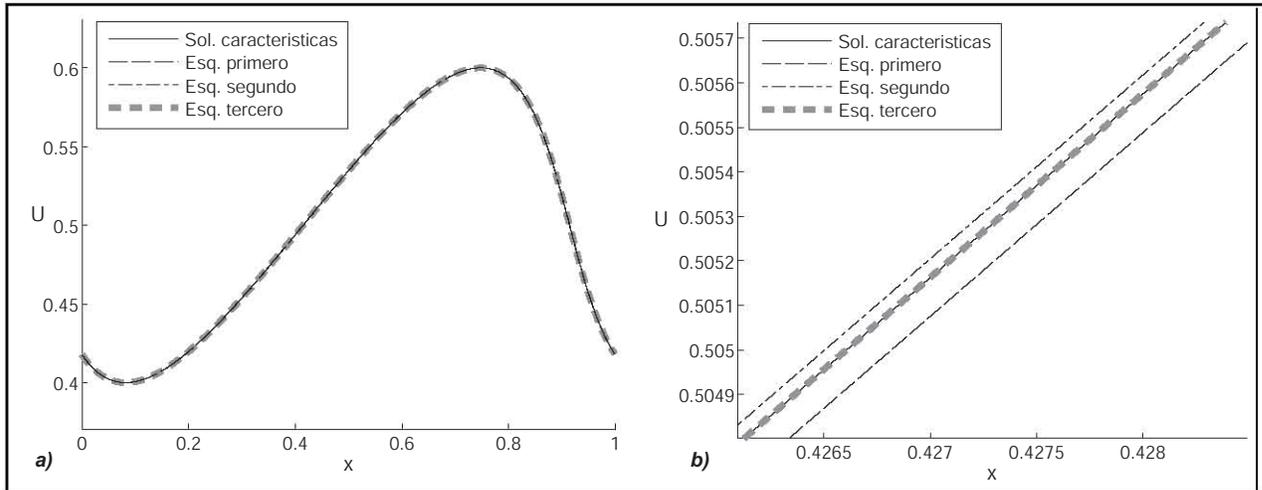


Figura 59: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales periódicas, $j = 830$, la figura **b)** es un detalle de la figura **a)**.

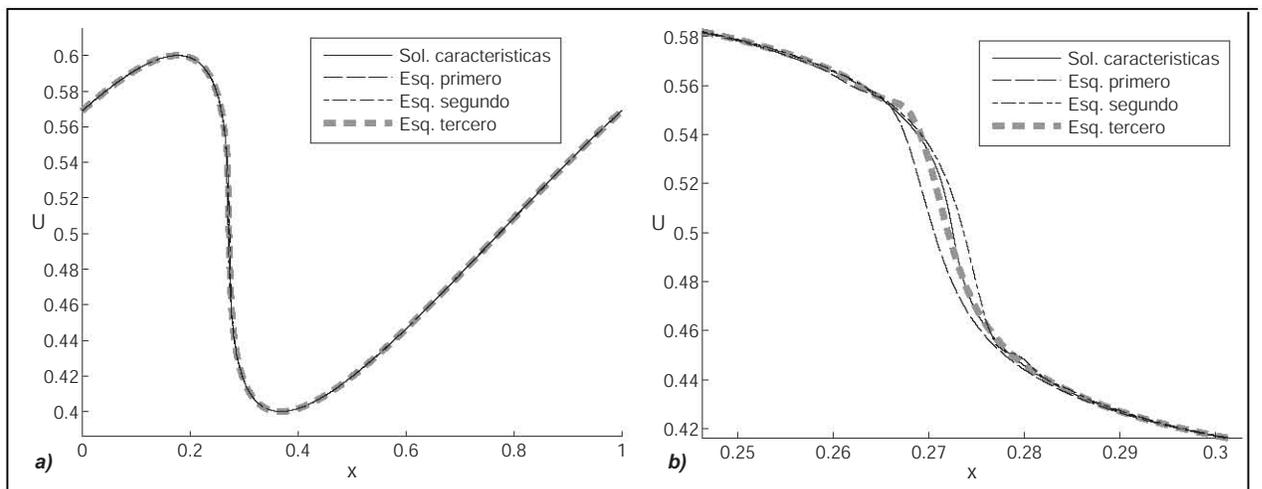


Figura 60: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales periódicas, $j = 1546$, la figura **b)** es un detalle de la figura **a)**.

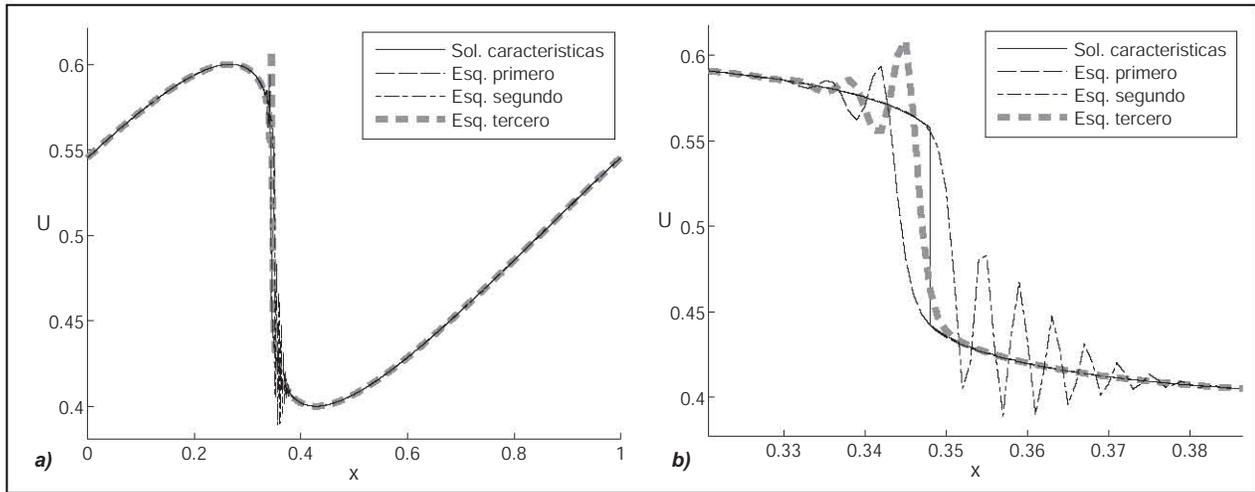


Figura 61: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales periódicas, $j = 1697$, la figura **b)** es un detalle de la figura **a)**.

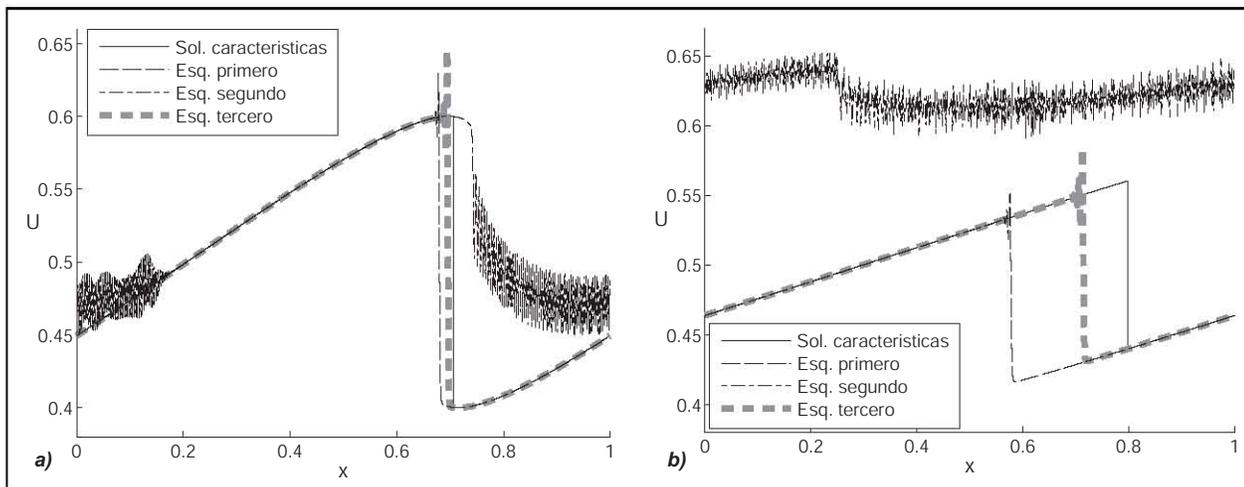


Figura 62: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales periódicas, **a)** $j = 2413$ y **b)** $j = 6596$.

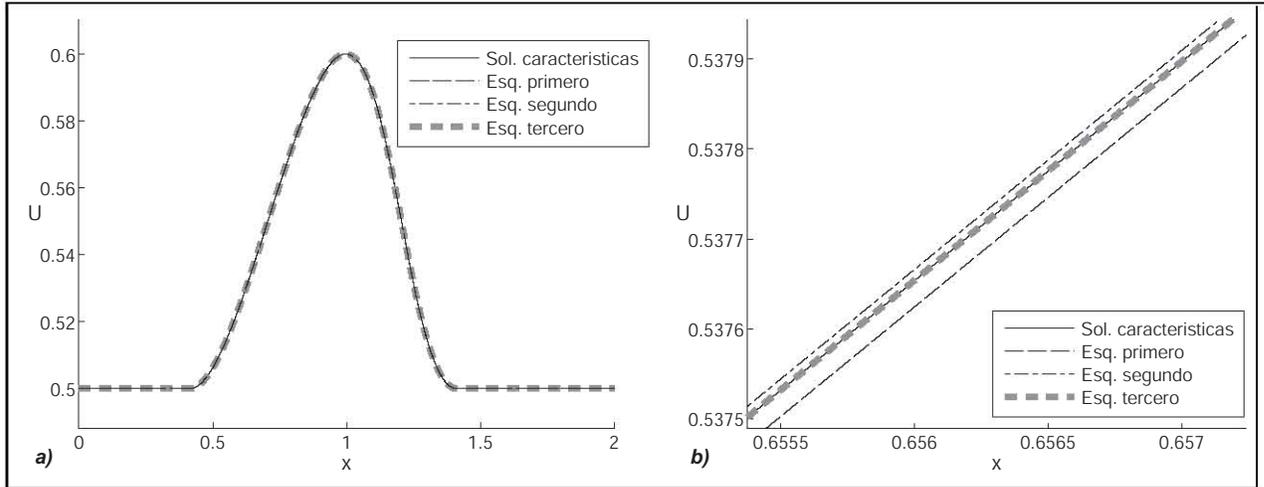


Figura 63: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales de soporte compacto, $j = 830$, la figura b) es un detalle de la figura a).

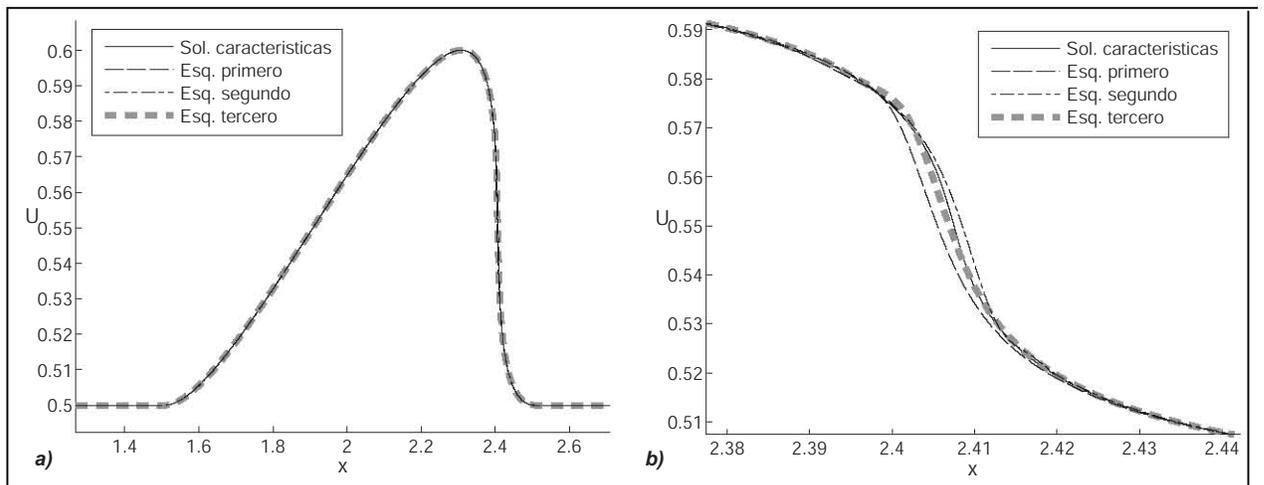


Figura 64: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales de soporte compacto, $j = 3016$, la figura b) es un detalle de la figura a).

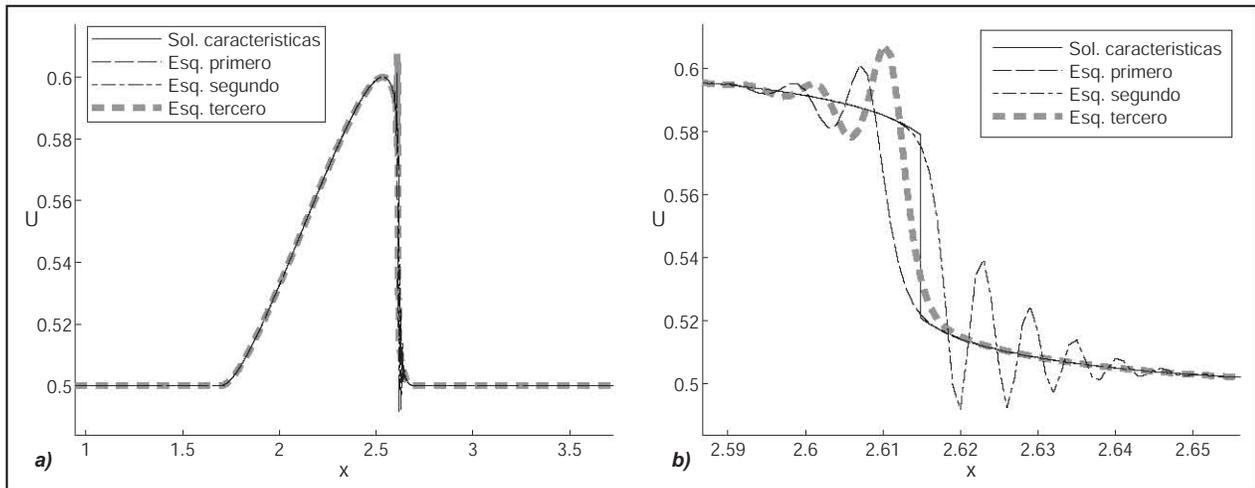


Figura 65: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales de soporte compacto, $j = 3393$, la figura b) es un detalle de la figura a).

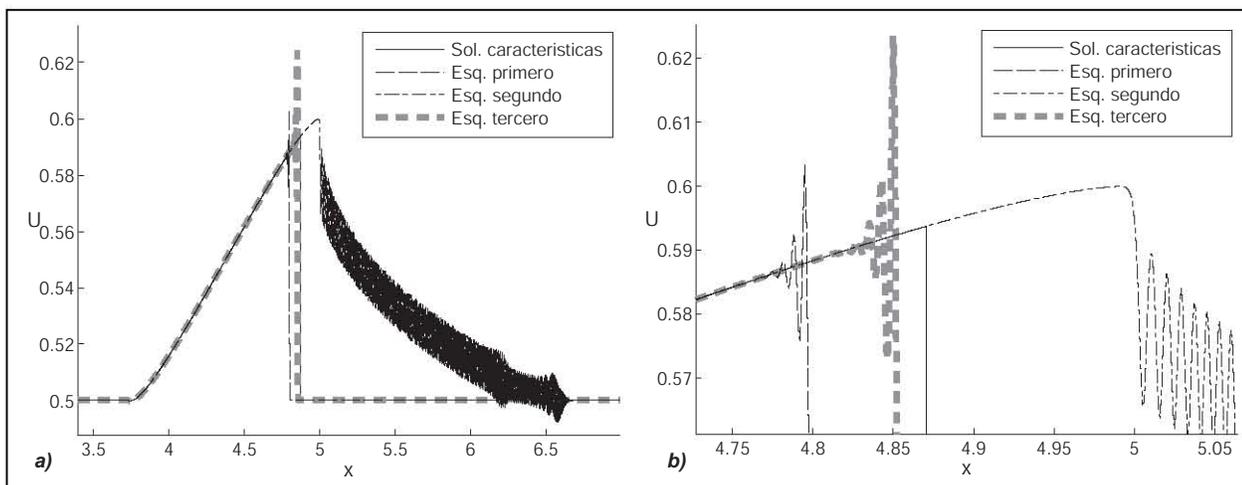


Figura 66: Resultados de esquemas numéricos relativos a la ecuación de Burgers no viscosa, para condiciones iniciales de soporte compacto, $j = 7500$, la figura b) es un detalle de la figura a).

8. Conclusiones y propuestas para trabajos posteriores

A partir de los resultados obtenidos en las secciones 5.6 y 7.5 concluimos que los esquemas numéricos que se han estudiado en el presente texto, siempre que se escojan adecuadamente sus parámetros, ofrecen resultados que aproximan bastante bien a la solución de la ecuación de advección e incluso la solución de la ecuación de Burgers no viscosa para tiempos anteriores a la formación de ondas de choque (discontinuidades). Sin embargo, de manera general presentan resultados no satisfactorios para los tiempos donde están presentes dichas discontinuidades: muestran oscilaciones que deterioran progresivamente la calidad de la aproximación de la solución numérica con respecto a la solución exacta. Atribuimos este hecho al carácter no lineal de la ecuación de Burgers no viscosa, caracterizada en parte por la eventual aparición de regiones donde crecen indefinidamente las pendientes hasta formar las mencionadas discontinuidades.

Sobre los esquemas empleados en el caso de la ecuación de advección, hemos observado que para un esquema numérico las propiedades que han resultado más importantes en cuanto a la calidad de sus resultados son: el orden del esquema numérico y el hecho de que sea explícito o implícito. Han ofrecido mejores resultados los esquemas de orden mayor (segundo orden en nuestro caso) y en el mismo sentido ha mostrado ser mejor el carácter implícito que el explícito. Es notable que en el caso de la ecuación de Burgers no viscosa, tales propiedades no han mostrado ser una ventaja significativa en cuanto a los resultados numéricos cerca de la discontinuidad; de hecho el deterioro de la aproximación ha sido mucho más pronunciado para el esquema implícito en comparación con el esquema explícito del mismo orden. Por otro lado, aunque el deterioro del esquema de segundo orden (esquema tercero) es menor que el de los otros dos esquemas, dicho deterioro no es cualitativamente diferente. Adicionalmente cabe señalar que, como puede observarse en los desarrollos expuestos en las secciones 5.5.3 y 7.4.2, la implementación de los esquemas implícitos ofrece dificultades adicionales, en particular cuando se trata de condiciones iniciales periódicas.

La otra manera de afectar la calidad de los resultados numéricos, como se ha mostrado en la sección 5.6, es mediante dos parámetros: i_p , correspondiente a la cantidad de puntos en cada nivel temporal de la malla, y C , correspondiente, siempre que c haya sido fijada (en nuestro caso fijada como $c = 1$), a la proporción que guardan Δx y Δt , es decir, al cociente $\frac{\Delta t}{\Delta x}$. En el caso de la ecuación de advección se ha verificado que, como podía esperarse, un valor más alto de i_p trae consigo mejores resultados numéricos. Para la misma ecuación en el caso de C se han hecho tres observaciones importantes: en primer lugar que el caso $C = 1$ reduce los esquemas numéricos a operaciones muy sencillas donde todos los resultados numéricos corresponden precisamente a la solución exacta (solución de D'Alembert); en segundo lugar, que el resto de los valores de C guarda una relación no monótona con respecto a la calidad de los resultados (lo cual es consistente con las observaciones hechas mediante ecuaciones modificadas en la sección 6.8 en los casos aplicables) y que para tener una idea más clara de dicha relación deberían hacerse propagaciones con diferentes valores de C , en particular dentro del intervalo $(0.5, 1)$; y en tercer lugar, que los esquemas contracorriente, de Lax-Friedrichs y de Lax-Wendroff son inestables para $C > 1$. Nuevamente resalta el hecho de que los resultados numéricos obtenidos para los esquemas de la ecuación de Burgers no viscosa no han sido satisfactorios, aún cuando en dichas propagaciones se han usado los valores de i_p y $\frac{\Delta t}{\Delta x}$ que en el caso de la ecuación de advección han ofrecido mejores resultados, que son: $i_p = 1000$ y $\frac{\Delta t}{\Delta x} = 1$.

También de los resultados expuestos en la sección 5.6 se observa, de acuerdo a los conceptos descritos en la sección 4.14, que los esquemas numéricos relativos a la ecuación de advección presentan tanto difusión numérica como dispersión numérica, exceptuando el caso del esquema de Wendroff que no presenta difusión numérica, lo cual además deriva del análisis teórico expuesto en la sección 5.4.

Si bien no es posible abarcar la totalidad de los esquemas de diferencias finitas para poder determinar propiedades generales de todos ellos, en este texto se han estudiado algunos esquemas que han sido escogidos buscando cubrir un espectro amplio de esquemas, y los resultados obtenidos han mostrado notables similitudes. Esto nos indica que podemos esperar que los resultados numéricos que pudieran obtenerse de otros esquemas numéricos de diferencias finitas, resulten también similares a los resultados que se han expuesto en el presente texto.

Otro aspecto relevante de los esquemas numéricos que se han estudiado es la importancia de su carácter lineal. Gracias a que los esquemas numéricos relativos a la ecuación de advección son lineales (debido a la linealidad de la ecuación de advección) fue posible aplicar una serie de herramientas teóricas sobre ellos para determinar principalmente dos cosas: primero, que es posible aplicar sobre ellos la prueba de estabilidad de Von Neumann para luego, usando el teorema de equivalencia de Lax, determinar la convergencia de dichos esquemas; y segundo, que es posible analizar las propiedades de los resultados que se obtendrían de cualquier condición inicial de soporte compacto, analizando los resultados que se obtienen a partir de condiciones iniciales del tipo senoidal, sumando posteriormente dichos resultados de modo apropiado, como se ha detallado en la sección 4.12.3. Ninguna de estas dos herramientas se ha podido aplicar en el caso de los esquemas relativos a la ecuación de Burgers no viscosa, y ello se debe a que dichos esquemas son no lineales, lo que a su vez deriva de que la ecuación de Burgers no viscosa es no lineal. En otras palabras, los esquemas relativos a la ecuación de Burgers no viscosa han sido implementados sin la certeza de que sean estables y en ninguno de estos casos podemos establecer analíticamente que los resultados obtenidos sean la suma de los resultados que se obtendrían de propagaciones con condiciones iniciales periódicas o de ningún otro tipo. En estos casos en general no es posible establecer estabilidad y convergencia de los esquemas analíticamente, y lo que se procura entonces es sencillamente variar parámetros como i_p y C de manera que pueda uno suponer razonablemente que sí se cuenta con estabilidad y convergencia. En problemas particulares a menudo se compara con resultados asintóticos para evaluar dichos esquemas.

Por otro lado, los resultados expuestos en la sección 6.8 nos muestran que las ecuaciones modificadas, hablando en términos generales, son una buena manera de describir los resultados numéricos que se obtendrían del esquema numérico al cual corresponden. En lo específico, dichos resultados establecen que los esquemas a los cuales corresponde una ecuación modificada del tipo advección-difusión presentan principalmente difusión numérica, y los esquemas a los cuales corresponde una ecuación modificada del tipo advección con dispersión presentan principalmente dispersión numérica. Otro aspecto importante con respecto a las ecuaciones modificadas, es que éstas finalmente también son ecuaciones diferenciales que pueden surgir en diferentes problemas relevantes en términos físicos y cuya solución puede querer aproximarse mediante esquemas numéricos de diferencias finitas. En dicho caso se pueden usar los esquemas contracorriente y de Lax-Friedrichs para la ecuación de advección-difusión, y los esquemas de Lax-Wendroff y de Wendroff para la ecuación de advección con dispersión. En ambos casos la aproximación es muy buena debido a que los dos primeros esquemas son de segundo orden con respecto a la ecuación de advección-difusión y los dos últimos son de tercer orden con respecto a la ecuación de advección con dispersión. Existe una limitación para el empleo de dichos esquemas debido al hecho de que, en las ecuaciones modificadas, tanto α como μ son funciones de Δx , Δt y c , sin embargo, es posible salvarla en buena medida por medio de un reescalamiento de las variables, es decir, del mismo modo que para los esquemas correspondientes a la ecuación de advección C juega un papel importante por que determina los casos que son equivalentes (que son aquellos donde el valor de C es el mismo, vea sección 4.5 y 5.5), así para el uso de dichos esquemas como aproximación de la ecuación de advección-difusión o de advección con dispersión, un cambio de variable del tipo $x' = k_1 x$, $t' = k_2 t$ (donde k_1 y k_2 son constantes) permite establecer una equivalencia entre determinados juegos de valores Δx , Δt , c y α (o bien μ),

y eso permite abarcar un rango amplio de casos.

Como se ha mencionado en la sección 6.8 las soluciones analíticas no necesariamente son más fáciles de implementar que los esquemas numéricos de diferencias finitas y esto se debe, por ejemplo, a que dichas soluciones analíticas pueden estar expresadas como sumas de un gran número de términos. Sin embargo, para fines de estimación de resultados por medio de las ecuaciones modificadas las soluciones analíticas de estas últimas tienen una ventaja: si se desea estimar el resultado que se obtendría por medio de un determinado esquema numérico para un valor de $j = j_0$ muy grande, puede implementarse la solución analítica de su ecuación modificada directamente en j_0 , sin necesidad de hacer cálculos para ninguno de los tiempos anteriores. Este no es el caso de los resultados obtenidos numéricamente, donde para obtener los valores correspondientes a j_0 se deben calcular los valores correspondientes a todos los tiempos anteriores.

La mayoría de las conclusiones señaladas pueden encontrarse en libros de texto de nivel introductorio sobre el tratamiento numérico de ecuaciones diferenciales. El autor considera que la aportación del presente texto constaría principalmente de dos partes: por un lado el uso que se le pueda dar al texto como introducción para estudiante de ciencias físicas y matemáticas a los aspectos básicos del área; y por otro lado la ilustración que se hace de las limitaciones en el uso de distintos esquemas de diferencias finitas para ecuaciones lineales y no lineales. De especial interés resulta el caso de la ecuación de Burgers no viscosa debido a que, si bien ningún texto recomienda el uso de esquemas de diferencias finitas en casos no lineales, tampoco es fácil encontrar en la literatura resultados numéricos como los que se obtuvieron para el presente texto, que muestren las limitaciones de dichos esquemas.

Por último mencionaremos dos ideas, que surgen como resultado del estudio realizado, y podrían tomarse en cuenta para buscar esquemas numéricos que brindaran resultados significativamente mejores que los esquemas expuestos:

- Un esquema numérico que supere las limitaciones que se han expuesto para los esquemas estudiados debería tener un comportamiento sustancialmente mejor en las zonas donde la solución exacta presente discontinuidades, por lo tanto una manera de buscarlo, sería evaluar los resultados que ofrece cuando se le implementa con condiciones iniciales discontinuas. Es decir, para evaluar la calidad de los resultados de cualquier otro esquema numérico que fuera propuesto, podría usarse, además de los dos tipos de condiciones iniciales que se han usado en el presente texto, una condición inicial sencilla pero discontinua, como por ejemplo, la correspondiente al problema de Riemann.
- De acuerdo al modo en que se ha construido la solución analítica de la ecuación no lineal estudiada en el presente texto, la ecuación de Burgers no viscosa, dicha solución consta de dos partes: una primera parte correspondiente a las regiones suaves, y una segunda asociada a la condición de Rankine-Hugoniot que debe satisfacerse en las discontinuidades. Este tipo de solución sugiere una alternativa, en lo que respecta a esquemas numéricos, que no corresponde a ninguno de los esquemas expuestos: podrían construirse esquemas numéricos que constaran también de dos partes, la primera parte para las regiones suaves, que podría estar constituida por un esquema numérico sencillo de diferencias finitas (que como se ha visto funciona bien en este tipo de regiones), y la segunda parte sería una restricción numérica que correspondiera a las condiciones de Rankine-Hugoniot que deben ser satisfechas precisamente en las discontinuidades.

Apéndices

A. Material digital interactivo.

Con el fin de ampliar las posibilidades de acceso para el lector a los resultados obtenidos de las diferentes implementaciones que se han expuesto, se anexa al texto un CD con una pequeña aplicación gráfica a la cual llamamos *material digital interactivo*.

Instalación.

Para poder usar este material, es necesario llevar a cabo un procedimiento de instalación que depende de la plataforma que se use:

Linux x86: Copie a una carpeta local el directorio `/Linuxx86/` que se encuentra dentro del CD, llamaremos `<RutaMDILocal>` a la ruta donde ha depositado la carpeta mencionada. Luego por medio de una terminal ubíquese en el directorio `<RutaMDILocal>`, una manera de hacerlo es ejecutando el comando

```
cd <RutaMDILocal>
```

en dicho comando debe sustituirse `<RutaMDILocal>` por la ruta que corresponde. Luego para abrir la aplicación ejecute el comando

```
./Menu.sh
```

Aunque el archivo `Menu` también es ejecutable no se recomienda usarlo porque muy probablemente será obtenido un error por falta de librerías. Un problema que puede presentarse es que cuando haga la copia mencionada a un directorio local, se pierdan los permisos de ejecución, si encuentra este tipo de problema puede resolverlo ejecutando, por ejemplo, el siguiente comando:

```
chmod -R a+x <RutaMDILocal>
```

donde, como antes, es necesario sustituir `<RutaMDILocal>` por la ruta correspondiente.

MS. Win: Copie a una carpeta local el directorio `\W32\`, que se encuentra dentro del CD, ejecute el instalador `MCRInstaller.exe` (el de su copia local) que se encuentra dentro de la carpeta `\InstaladorRuntime\` y siga las instrucciones que aparecen. Finalmente para acceder al material ejecute `Menu.exe` (también de su copia local).

Uso del material digital interactivo.

Al acceder al material debe aparecer una ventana de *Menu* como la que se muestra en la figura 67. Esta ventana contiene una breve descripción del material y cuenta con tres botones.

Al oprimir el primer botón de la ventana *Menu* debe aparecer una nueva ventana como la que se muestra en la figura 68, que permite tener acceso a las gráficas de propagación correspondientes a los esquemas numéricos de la ecuación de advección, así como a los resultados obtenidos al resolver sus respectivas ecuaciones modificadas. En la ventana hay una serie de controles, numerados en la misma figura 68, que se describen a continuación.

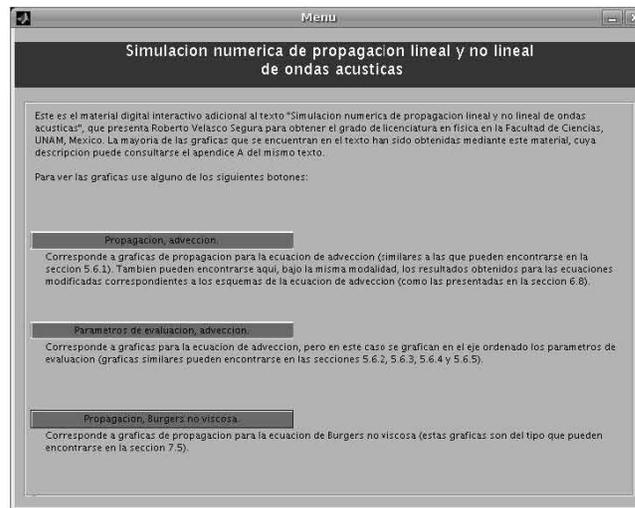


Figura 67: Ventana *Menu*.

- A.1** Selector de condiciones iniciales: permite seleccionar el tipo de condiciones iniciales que se han usado para la propagación numérica, ya sea condiciones iniciales periódicas, o bien condiciones iniciales de soporte compacto.
- A.2** Cuadros de selección de series graficadas: permiten seleccionar cuales series son graficadas. Se ha incluido este control porque frecuentemente sucede que las series de datos se tapan unas a otras, de modo que con este control puede desactivar momentánea o permanentemente una serie para poder observar las que están debajo.
- A.3** Controles de ejes: estos controles permiten indicarle a la aplicación los intervalos que deben mostrarse tanto en el eje de las abscisas (inhabilitado para el caso periódico), como en el de las ordenadas. El caso inhabilitado se debe a que para condiciones iniciales periódicas los resultados numéricos siempre están en el intervalo $[0, 1]$. También aparecen dos casillas de verificación; si están palomeadas se usan los valores introducidos, en caso contrario la aplicación usa un intervalo donde aparezcan todos los datos disponibles (vea más adelante **Resumen de datos**).
- A.4** Barra deslizable: permite indicar a qué valor de j (valor de t) deben corresponder las gráficas que se muestran. En la parte inferior aparecen los valores de Δx y Δt que corresponden a los datos mostrados.
- A.5** Selectores de los parámetros i_p y C : al oprimir sobre cualquiera de los valores disponibles se actualizará la gráfica para mostrar los resultados numéricos correspondientes a dicho valor. Este control está vinculado con el cuadro de selección A.6 que está justamente arriba. La opción por defecto es j de modo que al cambiar alguno de los parámetros se mostrarán nuevas gráficas correspondientes al valor actual de j ; si se cambia la opción a t , y se cambia por ejemplo el valor de C , la aplicación buscará (y mostrará las gráficas correspondientes) el valor de j cuyo valor correspondiente de t (para la nueva C) sea lo más próximo posible (vea más adelante **Resumen de datos**) al valor de t presente antes de cambiar el valor de C . Si el valor de j que correspondería sobrepasa $j = 2500$ se mostrará un mensaje de error.
- A.6** Selector de parámetro respetado al cambiar el valor de i_p o C : vea el inciso anterior.

Al oprimir el segundo botón de la ventana *Menu* debe aparecer una nueva ventana, como la que se muestra en

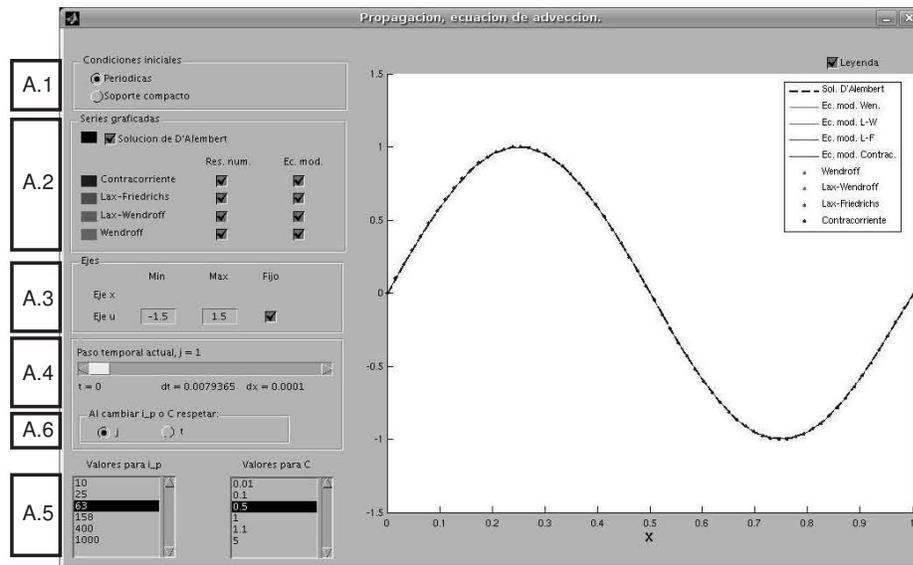


Figura 68: Ventana *Propagación, ecuación de advección*.

la figura 69, que permite observar gráficas donde el eje de las ordenadas es ocupado por alguno de los parámetros de evaluación (A amplitud, D desplazamiento o N norma). Este tipo de gráficas son similares las que aparecen en las secciones 5.6.2, 5.6.3, 5.6.4 y 5.6.5. En esta ventana hay controles que deben usarse antes de poder observar una gráfica, los cuales se detallan a continuación.

- B.1** Selector de parámetro para el eje ordenado: permite seleccionar el parámetro de evaluación que corresponderá al eje ordenado en las gráficas mostradas.
- B.2** Selector de condiciones iniciales: permite seleccionar el tipo de condiciones iniciales que se han usado para las propagaciones numéricas sobre los cuales se han medido los parámetros de evaluación que se muestran en las gráficas.
- B.3** Selectores para eje de las abscisas y series: hay cuatro parámetros que puede usarse en las gráficas para el eje de las abscisas o las series: i_p , C , j , “esquemas numéricos y ecuaciones modificadas”. Se debe escoger solamente uno de ellos para el eje de las abscisas y solamente uno para las series; cuando escoja el que corresponde a las series verá como cambia la zona C.3. Se recomienda (por claridad de la gráfica) que al empezar a usar la aplicación se use el parámetro j para el eje de las abscisas.
- B.4** Botón graficar: Una vez usados correctamente los tres controles anteriores oprima este botón y si todo está en orden el texto interior del botón debe decir “Cambiar parámetros” y debe aparecer una gráfica.

En la ventana *Parámetros de control y evaluación, ecuación de advección*, una vez obtenida la gráfica se deshabilitan algunos controles y se habilitan otros, como lo muestra la figura 70. Los controles habilitados se detallan a continuación.

- C.1** Controles de ejes: permiten establecer los rangos para los ejes de abscisas y ordenadas que han de mostrarse en las gráficas. Cuentan con una opción para uso manual o automático; en el caso automático la aplicación muestra un intervalo tal que todos los datos graficados queden dentro. También cuentan con una opción para escala lineal o logarítmica, que resulta útil para apreciar si un decaimiento es exponencial.

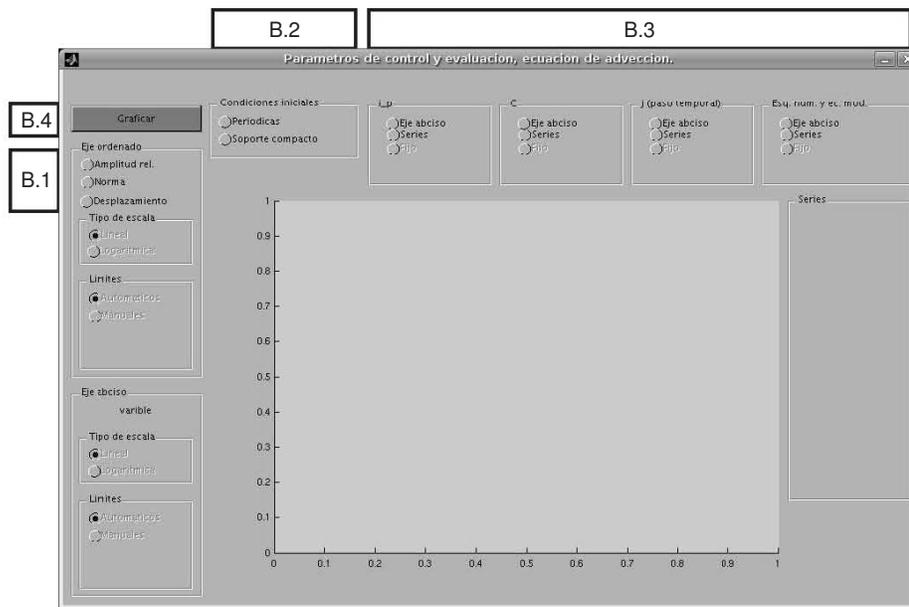


Figura 69: Ventana *Parámetros de control y evaluación, ecuación de advección*, primer estado.

C.2 Selectores de parámetros: para aquellos parámetros que no hayan sido escogidos para el eje del abscisas ni para las series, aparece un control que permite determinar el valor (de dichos parámetros) al cual deben corresponder las gráficas. En el caso de C , i_p y “esquemas numéricos y ecuaciones modificadas” aparece un menú desplegable y en el caso de j aparece un cuadro donde es posible escribir un valor. En este caso, al oprimir ENTER, la aplicación mostrará las gráficas para el valor disponible de j que sea más cercano (vea más adelante **Resumen de datos**).

C.3 Casilla de verificación para series: permiten determinar cuales series deben aparecer en las gráficas. Este control es útil porque frecuentemente suceden dos cosas: o bien las gráficas se tapan unas a otras, o bien los datos correspondientes a diferentes series se encuentran en órdenes de magnitud muy diferentes. En particular cuando las series corresponden al parámetro C , como algunos esquemas son inestables, los resultados numéricos pueden contener valores muy altos, por lo que se recomienda deshabilitar las series correspondientes a $C = 1.1$ y $C = 5$, usando los ejes en modo automático (ver inciso C.1).

C.4 Botón para cambio de parámetros: este botón regresa la ventana a su estado inicial (sin gráfica), vuelve a habilitar los controles B y deshabilita los controles C.

Cuando se oprime el tercer botón de la ventana *Menu* aparece una nueva ventana, donde se pueden observar los resultados numéricos obtenidos para los esquemas numéricos correspondientes a la ecuación de Burgers no viscosa. Los controles disponibles se describen a continuación.

D.1 Selector de condiciones iniciales: permite seleccionar el tipo de condiciones iniciales (periódicas o de soporte compacto) a las cuales corresponden los resultados numéricos que se muestran.

D.2 Casillas de selección de series graficadas: estas casillas permiten elegir los resultados numéricos a los cuales corresponden las series graficadas. Son útiles, por ejemplo, cuando las gráficas se tapan unas a otras.

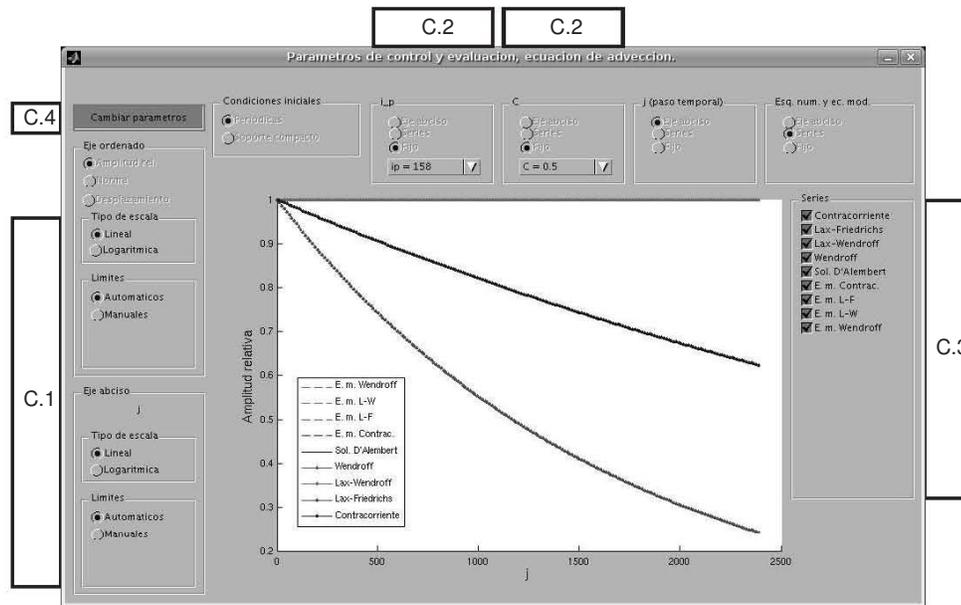


Figura 70: Ventana *Parámetros de control y evaluación, ecuación de advección*, segundo estado.

D.3 Controles de ejes: permiten determinar los intervalos de los ejes sobre los cuales han de mostrarse las gráficas. El control del eje de las abscisas está deshabilitado para el caso periódico porque los resultados numéricos siempre están dentro del intervalo $[0, 1]$.

D.4 Barra deslizante: permite determinar el valor de j (valor de t) al cual corresponden los resultados numéricos que se muestran. En la parte inferior se muestran los valores de t , i_p , Δx y Δt que corresponden a la gráfica mostrada.

Resumen de datos

Aún cuando en este *material digital interactivo* es posible observar una cantidad considerable de datos, no ha sido posible presentar la totalidad de ellos (que ocupan un espacio aproximado de 13GB); se ha tenido que hacer un *resumen*. En lugar de presentar los datos correspondientes a todos los valores de j se han escogido 200 valores distribuidos entre su mínimo ($j = 1$) y su máximo ($j = 2500$ para el caso de la ecuación de advección y $j = 7500$ para el caso de de la ecuación de Burgers no viscosa).

Del mismo modo, para la ecuación de advección, en el caso periódico, cuando $i_p > 200$ se han escogido 200 valores de i distribuidos uniformemente y son éstos los valores efectivamente graficados (para la escala que se maneja en la aplicación no se aprecia ninguna diferencia con respecto a las gráficas que incluyen todos los valores de i). En el caso de condiciones iniciales de soporte compacto, para al misma ecuación de advección, se grafican solamente los puntos que estén dentro de un intervalo ± 2 (en unidades de x) escogido alrededor del máximo de U para cada serie.

En el caso de los resultados numéricos correspondientes a la ecuación de Burgers no viscosa, sí se grafican todos los valores existentes de i , es decir, el eje x contiene todos los valores numéricos que han sido calculados, y esto se debe a que en la zona de discontinuidad es importante poder ver como cambia el valor de U con respecto a x ,

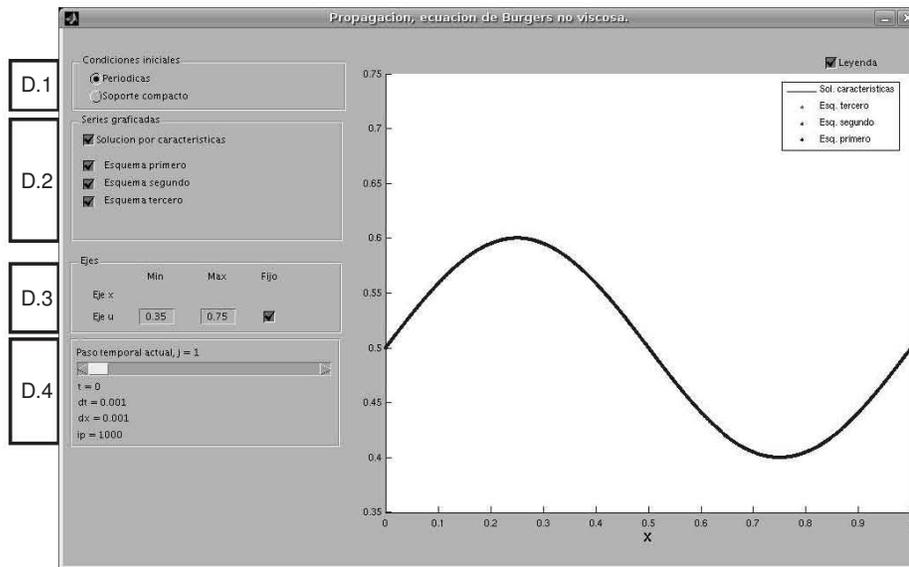


Figura 71: Ventana *Propagación, ecuación de Burgers no viscosa*.

de un punto a otro. Por la cantidad de datos en cada gráfica, es apreciable el hecho de que la aplicación responde notablemente más rápido para las ventanas correspondientes a la ecuación de advección.

Referencias

- [1] William F. Ames. *Nonlinear Partial Differential Equations in Engineering*. Academic Press, New York, 1965.
- [2] William F. Ames. *Numerical Methods for Partial Differential Equations*. Academic Press / Thomas Nelson & Sons, San Diego, 1977.
- [3] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists*. Harcourt Academic Press, San Diego, 5th edition, 2001.
- [4] G. R. Baldock and T. Bridgeman. *Mathematical Theory of Wave Motion*. Ellis Horwood Limited / John Wiley & Sons, New York, 1981.
- [5] J. Billingham and A. C. King. *Wave Motion*. Cambridge University Press, 2000.
- [6] William E. Boyce and Richard C. DiPrima. *Ecuaciones diferenciales y problemas con valores a la frontera*. Limusa Wiley, México, 4th edition, 2000.
- [7] Herbert B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. Wiley, Nueva York, 2nd edition, 1985.
- [8] R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves*. Springer, New York, 1948.
- [9] Richard Courant and Fritz John. *Introducción al cálculo y al análisis matemático Vol. 2*. Limusa, México, 2000.
- [10] Richard Courant and Fritz John. *Introducción al cálculo y al análisis matemático Vol. 1*. Limusa, México, 2001.
- [11] Frank S. Crawford. *Ondas, Berkeley Physics Course, Volumen 3*. Reverte, Barcelona, 1968.
- [12] Lokenath Debnath. *Nonlinear Partial Differential Equations for Scientist and Engineers*. Birkhäuser, Boston, 1997.
- [13] R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris. *Solitons and Nonlinear Wave Equations*. Academic Press, London, 1982.
- [14] DuChateau and Zachmann. *Applied Partial Differential Equations*. Dover, 2002.
- [15] Feynman, Leighton, and Sands. *The Feynman Lectures on Physics, Mainly Mechanics, Radiation, and Heat, Volume 1*. Addison-Wesley, 1963.
- [16] Eugene Isaacson and Herbert Bishop Keller. *Analysis of Numerical Methods*. John Wiley, 1966.
- [17] Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, and James V. Anders. *Fundamentals of Acoustics*. John Wiley & Sons, New York, 4th edition, 2000.
- [18] Kundu and Cohen. *Fluid Mechanics*. 2004.
- [19] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. Pergamon Press, 2nd edition, 1982.
- [20] Serge Lang. *Undergraduate Analysis*. Springer, New York, 1997.
- [21] Sidney Leibovich and Richard Seebass, editors. *Nonlinear Waves*. Cornell University Press, Ithaca, 1974.
- [22] Jerold E. Marsden and Michael J. Hoffman. *Elementary Classical Analysis*. W. H. Freeman and Company, New York, 1993.

- [23] Mathews and Fink. *Métodos Numéricos con MATLAB*. Prentice Hall, 3rd edition, 1999.
- [24] E. L. Miller. Predictor-corrector studies of burgers' model of turbulent flow. Master's thesis, University of Delaware, Newark, Delaware, 1966.
- [25] A. R. Mitchell and D. F. Griffiths. *The Finite Difference Method in Partial Differential Equations*. John Wiley & Sons, Chichester, 1980.
- [26] C. L. Mofrey and G. P. Howell. Nonlinear propagation of aircraft noise in the atmosphere. *AIAA Journal*, 19(8):986–992, August 1981.
- [27] K. Naugolnykh and L. Ostrovsky. *Nonlinear Wave Processes in Acoustics*. Cambridge University Press, 1998.
- [28] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C, The Art of Scientific Computing*. Cambridge University press, New York, 2nd edition, 1992.
- [29] M. I. Rabinovich and D. I. Trubetskov. *Oscillations and Waves in Linear and Nonlinear Systems*. Kluwer Academic Publishers, Dordrecht, 1989.
- [30] Robert D. Richtmyer and K. W. Morton. *Difference Methods for Initial Value Problems*. Interscience Publishers, New York, 2nd edition, 1967.
- [31] O. V. Rudenko. Interaction of intense noise waves. *Sov. Phys. Usp.*, 149(3):413–447, July 1986.
- [32] J.C. Strkwerda. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth & Brooks/Cole, 1989.
- [33] Eleuterio F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer, 2nd edition, 1999.
- [34] E. H. Twizell. *Computational Methods for Partial Differential Equations*. Ellis Horwood Limited / John Wiley & Sons, New York, 1984.
- [35] Randall J. Le Veque. *Numerical Methods for Conservation Laws*. Birkhauser Verlag, Basel, 1992.
- [36] G. B. Whitham. *Linear and Nonlinear waves*. John Wiley & Sons, New York, 1974.

